GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Integrating field and optical RapidEye data for above-ground biomass estimation: A study in the tropical peat-swamp forest of Sebangau, Central Kalimantan, Indonesia

Dissertation

to obtain the Ph.D. degree at the Faculty of Forest Science and Forest Ecology,
Georg-August-Universität Göttingen

submitted by

Damayanti Sarodja

Born in Jakarta, Indonesia

Göttingen, 2019

First Referee	:	Prof. Dr. Christoph Kleinn
Second Referee	:	Prof. Dr. I Nengah Surati Jaya
Examiner	:	Prof. Dr. Daniela Sauer
Date of Oral examination	:	20 December 2018

Acknowledgement

First of all, I would like to acknowledge my supervisor Prof. Dr. Christoph Kleinn for giving me the opportunity to conduct the research. I appreciate his excellent guidance, advices, and valuable comments throughout the research work. Many thanks to Dr. Lutz Fehrmann, Dr. Hans Fuchs, Dr. Paul Magdon, and Dr. Cesar Pérez-Cruzado for sharing the knowledge and for the fruitful discussions during my research work.

I cordially thank my second supervisor Prof. Dr. I Nengah Surati Jaya for his kind supports and valuable comments. I thank Dr. Suwido Limin and the team at CIMTROP for their efforts in the challenging fieldwork.

Deepest thanks to my friends and colleagues at the Chair of Forest Inventory and Remote Sensing: Henning Aberle, Ramon Trucios, Collins Kukunda, Dian Melati, Edwine Purnama, Almut Niebuhr, Kira Urban, Wanda Graf, Dengkui Mo, Xiaolu Tang, Nils Nölke, Philip Beckschäfer, Zihui Zhu, Sabine Schreiner, Haijun Yang, Silvia Wagner, Reinhard Schlote, and Hendrik Heydecke for all the helps and memorable times during my stay in Göttingen. Many thanks to Mats Mahnken and Lennard Dohrmann for their supports in the field data collections. I thank Mary Mulligan for proof reading all the texts. I also thank my colleagues in The Agency for the Assessment and Application of Technology, Indonesia, for their supports during my study leave.

My heartfelt gratitude to my beloved family: my mother, my husband Bobby Berlianto, brothers and sisters, nephews and nieces, for their unfailing faith, support, and love. Most of all, thank you for always believing in me even when I don’t believe in myself. There is no words could express how grateful I am to be part of you.

For my father and his unconditional love, I know you are always here with me.…

Sincerely,

Yanti Sarodja

This research was mainly funded by the Deutsche Forschungsgemeinschaft (DFG) KL 894/17. The study was conducted with the support from field datasets collected within the framework of: (i) The Biodiversity Explanatory research project, funded by the DFG (Priority Program 1374), (ii) SP05 of The Future Okavango Project (TFO), Namibia, between University of Göttingen, Germany, Stellenbosch University of South Africa, and Polytechnic University of Namibia.

Summary

The global impacts of greenhouse gas emissions from deforestation and forest degradation on climate change have increased international concern. An estimated 296 Gt of carbon is stored in the above- and below-ground biomass of global forests, which is almost half of the total carbon stored in forests. When forests are burned, degraded, or deforested, the stored carbon is released back into the atmosphere as carbon dioxide (CO_2). Over the last 25 years, the global amount of carbon stored in forest biomass has been decreased at an annual rate of about 0.44 Gt $year^{-1}$, or equivalent to 1.6 Gt $year^{-1}$ of CO_2. The majority of this loss occurred in tropical forests, which are known to reserve approximately 40% of the global terrestrial carbon. This concern led to an agreement in the forestry sector to reduce carbon emissions from deforestation and forest degradation, and to conserve, enhance, and sustainably manage forest carbon stocks, referred to as the REDD+ scheme. The basic idea of the scheme is for developed countries to compensate forest-rich developing countries in return for preserving carbon in their forests, while simultaneously allowing for necessary economic development.

The implementation of the REDD+ scheme requires an outline of a system for measuring, reporting, and verifying progress and changes. Paramount to this system is the establishment of "business-as-usual" baselines, against which the succession of carbon emission reductions of a country can be measured and compared. Hence, information on the amount of forest biomass and forest carbon stock is essential. This information can be used to estimate the amount of carbon that is potentially emitted into the atmosphere when a forest is burned or deteriorated. Such information is also useful for the stakeholders of forest planning, management and protection, as well as policy formulation. Combining remote sensing and field data has been recognized to increase the effectiveness in gathering this information, compared to that from the field data alone. The integration of remote sensing and field data also allows making wall-to-wall above-ground biomass (AGB) mapping over large areas. However, to achieve sound results, the integration requires the compatibility of the two datasets.

This study took place in the lowland tropical peatland of Sebangau, Central Kalimantan Province, Indonesia. Peat in this area has been accumulating for millennia. The forestlands, including the underlying thick peat deposits, play an important role in terrestrial carbon storage. Even so, numerous pressures to the area for decades have caused a large part of these peatlands to be devastated and vulnerable to fire. Part of the remaining forests also burned during the forest fire events in Indonesia, contributing to more carbon emission to the atmosphere. This study aims to contribute to the methodological basis of the integration of

field inventory and optical RapidEye data for forest AGB estimation using the case of Sebangau tropical peat-swamp forest. Three sub-studies of related topics were conducted: (i) the effect of forest restricted visibility in the basal area estimates from angle count plot method; (ii) the integration between field inventory and optical RapidEye data for AGB estimation model; and (iii) the effect of different field plot sizes in the AGB estimation models derived from field and RapidEye data.

Fixed-Area Plots and Angle Count Plots are frequently used plot designs in forest inventories. Angle Count Plot is known to be an efficient method for basal area estimation due to rapid application in the field. However, the estimation is rather sensitive to miscounted trees. In this study, the impact of the visibility condition in the Sebangau forest on the basal area estimation from Angle Count Plots was analyzed. Based on field measurements and simulation studies on a 1 ha plot of complete tree measurements, this study determined the maximum distance of visibility of the forest and the suitable basal area factor necessary to employ under this visibility condition. This study found a maximum distance of visibility of 6.6 m and recommends using a basal area factor of 5 for implementing the angle count plot method in the Sebangau forest to reduce the visibility effects on the estimation. The effects of restricted visibility in the Sebangau forest were further analyzed using field inventory data. For comparison, a dataset from an open savanna forest of Ncaute, Namibia, with almost no visibility restriction, was also used. The analyses were conducted for both the fixed-area- and the angle-count plot methods. The results showed that there was a visibility issue for estimating basal area using angle count plot in the Sebangau forest, while it was not found in the Ncaute forest.

For the second sub-study, the field plots of the Sebangau peat-swamp forest were combined with the predictor variables derived from RapidEye data to build an AGB estimation model through stepwise multiple linear regression. For better understanding, this modelling approach was applied to two different inventory datasets from a temperate forest of Hainich, Germany. The relationships between field-observed AGB and predictor variables derived from RapidEye data were analyzed across three different inventory datasets representing two different forest types. Results in the study showed that in general, the Pearson correlation coefficients r of field-observed AGB estimates and RapidEye predictors were weak for each dataset. The resulting AGB models exhibited weak performance, given by some common indicators such as Adj R^2 and $RMSE_r$. The AGB model for the Hainich forest, which was derived from a dataset with larger plot size, showed a better performance than other AGB models with an Adj R^2 of 0.65 and a $RMSE_r$ of 10.26%. The AGB estimations with integration of RapidEye data

showed a higher relative efficiency, in terms of their variances, when compared to the one derived solely from the field-data. Field plots in this study were according to the probability sampling design; the variance between the observed and the predicted AGB from the model was estimated using the model-assisted estimator, as well as the mean AGB estimate of each study area.

In the third sub-study, the effects of different plot sizes in the resulting AGB model were analyzed using three different scenarios with the Hainich forest dataset of larger plot size. In the first scenario, the same predictors were continuously used to estimate the AGB of different field plot sizes. In the second and third scenario, different predictors were allowed to be selected in the AGB model. Results from the first scenario clearly showed a decreasing performance of the AGB model with decreasing plot size. Additionally, in the second and third scenarios a decreasing pattern in the model's predictive power ($RMSE_r$) was found by decreasing the plot size. This pattern was not found in the model Adj R^2. Instead, the results showed that the relationships between the variability of the field-based AGB estimates and the variability of the RapidEye predictor variables are more complex. This study also showed that the plot perimeter lengths were significantly correlated with the model $RMSE_r$, as well as the model relative maximum residuals. These correlations were slightly higher than those between the plot sizes and the model $RMSE_r$ and the model relative maximum residuals. However, the AGB estimations with integration of RapidEye data showed a higher relative efficiency of 3.4 when compared to the one derived from the field-data only.

Zusammenfassung

Die globalen Auswirkungen der Treibhausgasemissionen aus Abholzung und Schädigung von Wäldern auf den Klimawandel haben die internationale Besorgnis erhöht. Eine geschätzte Menge von 296 Gt Kohlenstoff ist in der ober- und unterirdischen Biomasse der globalen Wälder gespeichert. Das entspricht fast der Hälfte des gesamten in Wäldern gespeicherten Kohlenstoffs. Wenn Wälder verbrannt, abgebaut oder abgeholzt werden, wird der gespeicherte Kohlenstoff als Kohlenstoffdioxid (CO_2) wieder in die Atmosphäre freigesetzt. In den letzten 25 Jahren wurde die Menge des in Waldbiomasse gespeicherten Kohlenstoffs weltweit mit einer jährlichen Rate von etwa 0,44 Gt, oder entsprechend 1,6 Gt CO_2, reduziert. Der größte Teil der Verluste ereignete sich in den tropischen Wäldern, wobei diese etwa 40 % des globalen terrestrischen Kohlenstoffs speichern. Dieses Problem führte zu einer Einigung über ein Schema im Forstsektor zur Verringerung der Kohlenstoffemissionen aus Entwaldung und Walddegradation und zur Erhaltung, Verbesserung und nachhaltigen Bewirtschaftung der Kohlenstoffbestände in Wäldern, das als REDD+ bezeichnet wird. Die Grundidee dieses Systems besteht darin, dass die Industrieländer waldreiche Entwicklungsländer als Gegenleistung für den Erhalt des Kohlenstoffs in ihren Wäldern entschädigen und gleichzeitig die notwendige wirtschaftliche Entwicklung ermöglichen.

Die Umsetzung des REDD+-Schemas erfordert den Entwurf eines Systems zur Messung, Berichterstattung und Verifizierung, und auch eine Business-as-usual Baseline, woran die CO_2-Emissionsreduktionen eines Landes gemessen und verglichen werden kann. Daher sind Informationen über die Menge an Forstbiomasse und den Waldkohlenstoffbestand unerlässlich. Die Informationen können verwendet werden, um die Menge an Kohlenstoff zu ermitteln, die bei der Verbrennung oder Zerstörung des Waldes potenziell in die Atmosphäre abgegeben wird. Die Informationen sind aber auch nützlich für Interessengruppen der Forstplanung, -bewirtschaftung und -schutz sowie der Politikgestaltung. Die Kombination aus Fernerkundungs- und Felddaten ist anerkannt, um die Effizienz bei der Erfassung dieser Informationen zu erhöhen, verglichen mit derjenigen, die nur aus Felddaten gewonnen wird. Die Verknüpfung von Fernerkundungs- und Felddaten ermöglicht auch die großflächige Wall-to-Wall-Kartierung von oberirdischer Biomasse (above-ground biomass, AGB). Um jedoch gute Ergebnisse zu erzielen, erfordert die Verknüpfung die Kompatibilität beider Datensätze.

Diese Studie wurde im tropischen Tiefland-Torfgebiet in Sebangau in der Provinz Zentral-Kalimantan, Indonesien, durchgeführt. Der Torf in diesem Gebiet wird seit Jahrtausenden angebaut. Die Waldflächen einschließlich der darunter liegenden dicken Torfschichten spielen eine wichtige Rolle als terrestrische Kohlenstoffspeicher. Dennoch hat der jahrzehntelange

Druck auf das Gebiet dazu geführt, dass ein großer Teil dieser Torfmoore heute verheert und durch Brände gefährdet ist. Ein Teil der verbliebenen Wälder, die auch bei den Waldbränden in Indonesien zerstört wurden, trug zu mehr Kohlenstoffemissionen in die Atmosphäre bei. Diese Studie soll zur methodischen Grundlage der Kombination von Feldinventur und optischen RapidEye-Daten mit einer relativ hohen räumlichen Auflösung zur Wald-AGB-Schätzung am Beispiel des tropischen Torf-Sumpfwaldes von Sebangau beitragen. Es wurden drei Teilstudien zu angrenzenden Themen durchgeführt: (i) der Effekt der eingeschränkten Sichtbarkeit im Wald in den Grundflächenschätzungen aus Winkelzählprobeflächen; (ii) die Kombination aus Feldinventur und optischen RapidEye Daten für das AGB-Schätzmodell; und (iii) der Effekt verschiedener Plotgrößen in den AGB-Schätzmodellen, die aus Feld- und RapidEye Daten abgeleitet sind.

Probeflächen mit fester Größe und Winkelzählprobeflächen sind die am häufigsten verwendeten Plotdesigns in der Waldinventur. Die Winkelzählprobe ist bekannt dafür, dass sie aufgrund ihrer schnellen Anwendung im Feld eine effiziente Methode zur Schätzung der Grundfläche ist. Diese Schätzung ist jedoch eher sensibel gegenüber falsch gezählten Bäumen. In dieser Arbeit wird analysiert, wie sich die Sichtverhältnisse im Wald von Sebangau auf die Grundflächenschätzung aus der Winkelzählprobe auswirken. Basierend auf Feldmessungen und Simulationsstudien auf einem 1 ha großen Plot mit kompletten Baummessungen wurde in dieser Studie die maximale Sichtweite des Waldes und der geeignete Grundflächenfaktor ermittelt, der unter diesen Sichtverhältnissen anzuwenden ist. Diese Studie ermittelte eine maximale Sichtweite von 6,6 m und empfiehlt, einen Grundflächenfaktor von 5 für die Implementierung der Winkelzählprobenmethode im Sebangauer Wald zu verwenden, um den Effekt der Sichtverhältnisse auf die Schätzung zu reduzieren. Die Auswirkungen der eingeschränkten Sicht in Sebangau wurden anhand von Waldinventurdaten weiter analysiert. Zum Vergleich dazu wurde auch ein Datensatz aus einem offenen Savannenwald von Ncaute, Namibia, mit nahezu keiner Sichteinschränkung verwendet. Die Analysen wurden sowohl für den Ansatz der Festflächen als auch für den Ansatz der Winkelzählprobenplots durchgeführt. Die Ergebnisse zeigen, dass es ein Sichtbarkeitsproblem bei der Schätzung der Grundfläche mit der Winkelzählmethode im Sebangauer Wald gibt, während sie im Ncaute-Wald nicht vorkommt.

Für die zweite Teilstudie wurden die Feldplots des Sebangauer Torf-Sumpfwaldes mit den aus den RapidEye-Daten abgeleiteten Prädiktorvariablen kombiniert, um ein AGB-Schätzmodell durch schrittweise multiple lineare Regression aufzubauen. Zum besseren Verständnis wurden in der Studie auch zwei verschiedene Inventurdatensätze aus einem gemäßigten Wald im

Hainich, Deutschland, verwendet. Die Beziehungen zwischen der feldbeobachteten AGB und den aus den RapidEye-Daten abgeleiteten Prädiktorvariablen, sowie die Performance jedes AGB-Modells aus verschiedenen Waldtypen mit unterschiedlichen Datensätzen, wurden analysiert. Die Ergebnisse der Studie zeigen, dass im Allgemeinen die Pearson-Korrelationskoeffizienten *r* der feldbeobachteten AGB-Schätzungen und der RapidEye-Prädiktoren schwach sind für jeden Datensatz. Die daraus resultierende Performance der AGB-Modelle war ebenfalls schwach, was durch einige gängige Indikatoren wie Adj R^2 und *RMSEr* bestätigt ist. Das AGB-Modell für den Hainicher Wald, das aus einem Datensatz mit größerer Plotgröße abgeleitet wurde, zeigte eine bessere Performance als andere AGB-Modelle mit einem Adj R^2 von 0,65 und einem *RMSEr* von 10,26%. Die Feldplots in dieser Studie wurden nach dem Konzept der Wahrscheinlichkeitsstichprobe erstellt, und die modellgestützten Schätzer wurden zur Schätzung der AGB verwendet.

In der dritten Teilstudie wurden die Auswirkungen unterschiedlicher Plotgrößen im resultierenden AGB-Modell anhand von drei verschiedenen Szenarien analysiert. Im ersten Szenario wurde immer das gleiche Modell verwendet, um die AGB des gleichen Gebietes mit unterschiedlichen Feldplotgrößen zu schätzen. Im zweiten und dritten Szenario konnten verschiedene Prädiktoren im Modell verschiedener Plotgrößen ausgewählt werden. Mit Hilfe dieser Modelle verschiedener Prädiktoren wurde die AGB des Untersuchungsgebietes geschätzt. Die Ergebnisse des ersten Szenarios zeigten deutlich eine abnehmende Performance des AGB-Modells bei abnehmender Plotgröße. Aus dem zweiten und dritten Szenario wurde ein abnehmendes Muster der Modellvorhersageleistung ($RMSE_r$) durch die Verkleinerung der Plotgröße festgestellt. Dieses Muster wurde nicht im Modell-Adj R^2 gefunden. Vielmehr zeigten die Ergebnisse, dass die Zusammenhänge zwischen der Variabilität der feldbasierten AGB-Schätzungen und der Variabilität der RapidEye Prädiktorvariablen komplexer sind. Diese Arbeit zeigte auch, dass die Längen der Plotperimeter signifikant mit dem Modell $RMSE_r$ und mit den relativen maximalen Residuen des Modells korreliert sind. Diese signifikanten Korrelationen waren etwas höher als die zwischen den Plotgrößen und dem Modell- $RMSE_r$ und den relativen maximalen Residuen des Modells. Die AGB-Schätzungen mit Integration von RapidEye-Daten zeigten jedoch eine höhere relative Effizienz von 3,4 im Vergleich zu derjenigen, die nur aus den Felddaten abgeleitet wurde.

Table of Contents

List of Figures

List of Tables

Abbreviation

ACP	:	Angle Count Plot
AGB	:	Above-Ground Biomass
AOT	:	Aerosol Optical Thickness
BA	:	Basal Area
BAF	:	Basal Area Factor
BAU	:	Business as Usual
CIMTROP	:	Center for International Co-operation in Management of Tropical Peatland
CO_2	:	Carbon dioxide
COP	:	Conference on the Parties
DBH	:	Diameter at Breast Height
DOD	:	Diameter over Deformation
EF	:	Expansion Factor
FAO	:	Food and Agriculture Organization
FAP	:	Fixed Area Plot
FC	:	Full Census
FCC	:	Falls Color Composite
FI	:	Forest Inventory
FREL	:	Forest Reference Emission Level
GCP	:	Ground Control Point
GHG	:	Green House Gas
GLCM	:	Grey Level Co-occurrence Matrix
GPS	:	Global Positioning System
HIL	:	High Intensity Lidar
IPCC	:	Intergovernmental Panel on Climate Change
LiDAR	:	Light Detection and Ranging
LOOCV	:	Leave-One-Out Cross-Validation
LSI	:	Large Scale Inventory
NFI	:	National Forest Inventory
NIR	:	Near Infrared
NLPF	:	Natural Laboratory for the Study of Peat-Swamp Forest
QMD	:	Quadratic Mean Diameter
RE	:	Relative Efficiency
REDD+	:	Reducing Emissions from Deforestation and forest Degradation, and foster conservation, sustainable management of forests, and enhancement of forest carbon stock
SAR	:	Synthetic Aperture Radar
TOA	:	Top of Atmospheric
UNFCCC	:	United Nations Framework Convention on Climate Change
UTM	:	Universal Traverse Mercator
VI	:	Vegetation Index
VIF	:	Variance Inflation Factor
WD	:	Wood Density
WGS	:	World Geodetic System

Chapter 1

Introduction

1.1. Carbon emission from forests and peatlands

Forests cover about 31 percent of the total global land surface and, provide a large habitat for terrestrial biodiversity (FAO, 2010), ecosystem services for local and global communities, and protection of land and water resources (Krieger, 2001). Through the processes of evapotranspiration, respiration, and photosynthesis, forests play a crucial role in global water and carbon cycles (Bonan, 2008). By absorbing carbon dioxide (CO_2) from the atmosphere and storing this carbon, forests act as carbon sinks and contribute to the mitigation of climate change. The sequestered carbon is accumulated in the form of forest living biomass, dead wood, litter, and forest soil. An estimated amount of 296 Gt of carbon is stored in the above- and below-ground biomass of global forests, which is almost half of the total carbon stored in all forests worldwide (FAO, 2016).

When forests are burned, degraded, or deforested, the stored carbon is potentially released back into the atmosphere as CO_2. In this way, the forests are functioning as a source of carbon. The amount of global carbon stocks in forest biomass has decreased by nearly 11 Gt in the last 25 years. The decrease in forest carbon stocks was equivalent to 0.44 Gt $year^{-1}$, or about 1.6 Gt $year^{-1}$ of CO_2 (FAO, 2016). Among the main reasons for the decreased carbon stocks were the conversion of forest lands for other land uses, predominantly as agriculture fields and settlements, or exploitation for timber. In 2014, the Intergovernmental Panel on Climate Change (IPCC) indicated that the combined sectors of agriculture, forestry and other land uses are the second largest source of Green House Gas (GHG) emissions after energy production. Forestry and land uses alone were reported to account globally to nearly 12% of the anthropogenic CO_2 emissions over the period of 2000-2009 (IPCC, 2014).

1. Most of the global forest loss occurs in tropical forests. These losses are significant since tropical forests account for 40% of the total carbon stored in the terrestrial biosphere (Meister et al, 2012). During the period from 1990-2015, great losses of carbon stocks in forest biomass occurred in Africa, South America, and South and Southeast Asia, including Indonesia (Fig. 1.1). The Indonesian tropical forest is the

third largest tropical rainforest in the world (Baccini et al., 2017). The forests are rich in biodiversity and store an enormous amount of carbon. Despite their importance, Indonesian forests have been under significant pressures for decades. Forests decreased from 118.5 Mha in 1990 to 91.0 Mha in 2015, with an annual deforestation rate of 1.1 Mha year^{-1} over the period 1990-2015 (FAO, 2015). From 2000 to 2010, deforestation in Indonesia, including peatland deforestation, was motivated mainly by agricultural expansion, particularly of crop and tree-pulp plantation development, forest fire, and peat fire (Verchot et al., 2010). In the year 2014, the Land Use, Land Use Change, and Forestry (LULUCF) together with peat fire in Indonesia, accounted for 53.1% of the total carbon emission of the country (Masripatin, 2017).

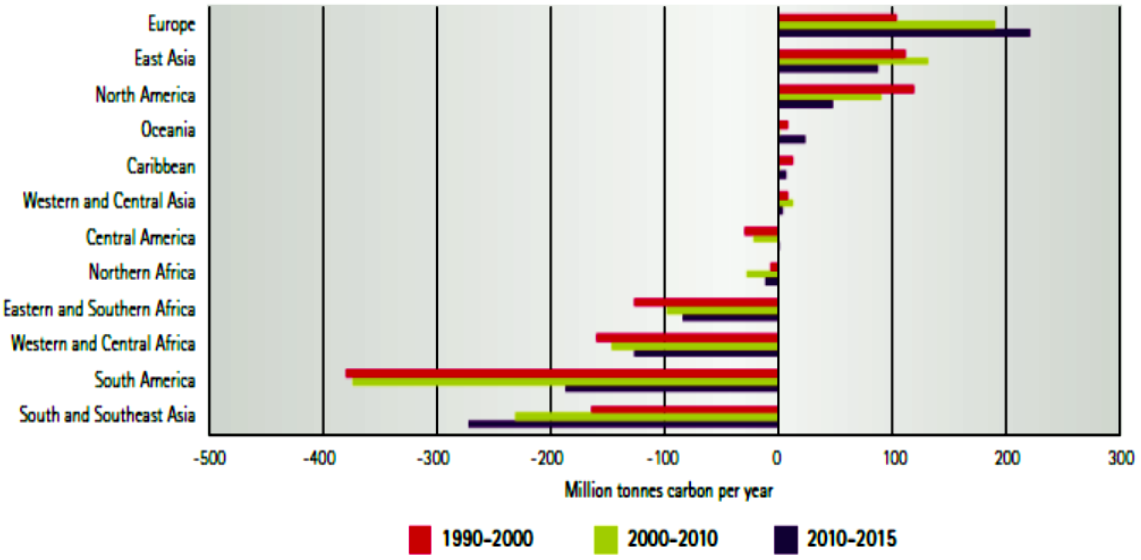


Fig. 1.1. Change of carbon stock in forest biomass 1990-2015 (FAO, 2016). Negative changes are net losses whereas positive changes are net gains.

Indonesia comprises the largest area of peat within Southeast Asia (Page et al., 2011). The peatlands contain an estimated total carbon stock of 74 Gt, out of which 19 Gt is stored aboveground in forests and 55 Gt belowground in peat deposits (Baccini et al., 2012; Jaenicke et al., 2008; Page et al., 2011). The peatlands are mostly covered by forests and are distributed predominantly in the three largest islands of Indonesia: Sumatera, Kalimantan, and Papua (Page et al., 2011). Essentially, peat is partially decayed organic material. In peatlands, the decaying of organic material is inhibited by acidic and anaerobic conditions. Since the annual rate of organic matter production is higher than the rate of decomposition, peat has been accumulating over thousands of years, although at a very slow pace. In general, peatlands

sequester a large portion of CO_2 and produce methane (CH_4) and nitrous oxide (NO_2), among other GHGs. Therefore, peatlands are either carbon sinks or sources, varying over time according to changing environmental conditions and the level of human disturbance (Strack, 2008). The carbon balances of peatlands may be altered by various circumstances including drainage, fire, and land clearance through deforestation for logging or agricultural conversion. When the peat is drained, the organic carbon is exposed to the air and turns into CO_2 by reacting with oxygen, which is released to the atmosphere (IPCC, 2014). The dry surface of the drained peat is then also susceptible to fire. In the last few decades, Indonesian forested peatlands have experienced frequent fires that have led to the reduction of their carbon stores (Hooijer et al., 2010, 2006; Page et al., 2002), e.g. fires in 1997, 2002, 2006, 2009, 2013, 2014, 2015 (Page and Hooijer, 2016). Peat fires cause the release of a significant amount of carbon to the atmosphere. The smoke from peat fires blankets large areas, causing social economic losses and directly affecting the lives of the people in their daily routine (Page and Hooijer, 2016). Regarding land clearance, removal of above-ground vegetation on peatland reduces the capacity of sequestration of the CO_2 released from the peat. Apart from their carbon-related functions, peatlands also play an essential role in the regulation of water flow, in the control of erosion and floods, in stabilizing micro- and macro-climates, and in providing forest resources and habitats for flora and fauna, some of which are endemic to these ecosystems (Strack, 2008).

1.2. International concern on carbon emissions from forestry sector

Growing awareness among international communities on the importance of forests in the global carbon cycle has increased concerns on the negative impact of GHG emissions from deforestation and forest degradation on climate change. Deforestation causes the loss of forest biomass as well as habitat and species loss and general environmental degradation. Disturbances on forests also negatively influence the capacity of carbon sequestration (Houghton, 1991).

In December 2007, the 13th Conference of the Parties (COP) of the United Nations Framework Convention on Climate Change (UNFCCC), held in Bali, came to an agreement within the forestry sector to reduce GHG emissions from deforestation and forest degradation and to conserve, enhance, and sustainably manage forest carbon stocks. This agreed upon scheme is referred to as the REDD+ activities. The basic idea of the scheme is for developed countries to compensate forest-rich developing countries in return for reducing their carbon emissions from forests (Cronin and Santoso, 2010). The December 2015 Paris COP encourages the

parties to continuously move towards the implementation of the REDD+ scheme, as written in the article 5.2 of the Paris Agreement, thereby undertaking ambitious efforts to combat climate change and to adapt to its effects. Furthermore, the Paris Agreement encourages developed countries to enhance support and assistance for developing countries to reach the stipulated targets. In this regard, all parties are required to put forward their climate actions plans intended for the period after 2020 into the so-called nationally determined contributions (NDCs) (article 4.2 of the Paris Agreement).

However, the implementation of the REDD+ scheme is not as simple as it sounds. It needs the outline of a system for measurement, reporting, and verification (MRV) as well as Business as Usual (BAU) baselines, against which reductions can be measured. In paragraph 71 of Decision 1/CP.16 of the Cancun agreement, the developing countries that aim to participate in the REDD+ are requested to develop the following elements: (i) REDD+ National Strategy or Action Plan, (ii) Forest Reference Emission Level/Forest Reference Level (FREL/FRL), (iii) A robust and transparent National Forest Monitoring System (NFMS), (iv) Safeguards Information System. To provide transparent, consistent, and reliable information, the UNFCCC encourages the Convention's Parties to use the guideline developed by the IPCC as the methodological basis for estimating and reporting the GHG emissions and removals of their country (UNFCCC, 2009). The approaches to obtain information on carbon stock changes were classified into three levels of details, so-called tiers. The accuracy of the estimates increases with the increase of each tier, as well as the data requirements and analytical complexity.

The development of the REDD+ elements mandates improvement in forest carbon stock estimation and forest resource monitoring methods. There are five carbon pools in which changes in ecosystem carbon stocks are to be monitored and reported: above-ground carbon (AGC), below-ground carbon (BGC), litter, dead wood, and soil (UNFCCC, 2009). The AGC, together with BGC, holds the most significant part of carbon at about 58%, while approximately 41% and 1% are stored in the soil and as litter, respectively (Meister et al., 2012).

1.3. REDD+ in Indonesia: International commitment to national actions

Indonesia is a continuous participant in international climate change mitigation activities, both through advocating on behalf of the global community and through being aware that the country itself is vulnerable to the adverse impacts of climate change (PEACE, 2007). The

country signed the UNFCCC at the Earth Summit held in Rio de Janeiro in 1992 and ratified the Convention in 1994 when it came into force. During the G-20 Pittsburg meeting on September 2009, the government of Indonesia pledged to reduce the country's GHG emission by 26% unconditionally under the BAU scenario by year 2020, or by 41% with international support.

Following the 2009 commitment, the government of Indonesia is progressively formulating the legal and policy instruments to reach the target. The same target was stipulated in the Presidential Regulation (PERPRES) No. 61/2011, on the National Action Plan on GHG Emissions Reduction by 2020, both unconditionally and with international support. The government also published the Presidential Regulation No. 71/2011 on the national GHG inventory guidance. To reduce the emissions from the land use sector, a moratorium on primary forest clearance and peatland conversion was first declared in 2011 under the Presidential Instruction (INPRES No.10/2011) and has since been renewed every two years (INPRES No. 6/2013, No. 8/2015, No. 6/2017). This moratorium policy aims to halt the new concession on primary forests and peatlands, with some exemptions on the following conditions: (i) if the permit of the new concession was already issued, (ii) if the area is required to support the national priorities program including geothermal, oil and gas, electricity, rice and sugarcane, (iii) if the ongoing concession extends their permit before it is terminated, and (iv) if the area is used for ecosystem restoration. Moreover, and due to the fact that palm oil is among the major causes for deforestation in Indonesia (Austin et al., 2017), the government recently signed a moratorium to halt the issuance of new permits for palm oil plantations and call for a review of existing palm oil plantations for the next three years, starting from September 19, 2018 (INPRES No. 8/2018). Another relevant Government Regulation (PP) was the PP No. 71/2014 on the protection and management of peat ecosystems and was revised in 2016 (PP No. 57/2016) to strengthen the effort for protecting and managing peatlands in a sustainable way, including restoration of degraded peatlands.

In January 2016, Indonesia submitted their forest reference emission level (FREL) for REDD+ to the UNFCCC Secretariat. The IPCC Guideline 2006 was used as general reference for measuring the emissions from deforestation, forest degradation, and peat decomposition. The FREL was calculated according to the tier 2 approach, where the country-specific emission factors and other parameters were applied using the reference period of 1990-2012. Based on the result, the Indonesian FREL from deforestation, forest degradation, and the associated peat decomposition was set at 0.568 Gt equivalent CO_2 $year^{-1}$ for the reference period of 1990-2012

(MoEF, 2016). The value was taken as the reference point against the actual carbon emission rate per year starting from 2013 to 2020 of the commitment period.

In the FREL document, some opportunities to improve the current estimates using the results from the on-going initiatives, previously unused in the development of FREL, were also mentioned. For example, to increase the availability of field data used in determining forest emission factors, one option is to include data from permanent research plots by Forest Research and Development Agency (FORDA) in addition to the national forest inventory (NFI) data. The permanent research plots were established since 2011 in 13 provinces. The plot design followed the Indonesian National Standard for ground-based forest carbon accounting (SNI 7724: 2011) with a square plot of 20 m x 20 m size. In future, they are expected to compliment 4450 NFI plots collected across the country within the period of 1990-2013 (MoEF, 2016). The Indonesian NFI was initiated by the Ministry of Forestry and started in 1989. Each sample in the Indonesian NFI design consisted of a cluster with nine sub-plots of 100 m x 100 m size, out of which only one sub-plot is permanent. The other sub-plots are temporary, where the angle count plot approach is implemented to collect data. This NFI project was supported by the Food and Agriculture Organization (FAO) of the United Nations and the World Bank and is now financed from the national budget. The other option is to improve the NFI data through enhancement of measurement precision and validation of current plots. Regarding improvement of the activity data, the potential use of higher resolution imagery other than Landsat will be explored. To achieve this, coordination with Indonesia's Aeronautics and Space Agency (LAPAN) is required for the provision of high resolution satellite imagery. The use of digital image classification, instead of the time-consuming visual interpretation methods, will also be examined.

Following up with the country's commitments, Indonesia submitted their first NDC to the UNFCCC in November 2016. In this NDC, Indonesia committed to unconditionally reducing the country's GHG emission by 29% below the BAU scenario by 2030, or by 41% below the BAU level with international support.

1.4. Estimating forest carbon/forest biomass

Biomass is commonly used as a proxy to estimate the amount of carbon preserved in forests (Meister et al., 2012), assuming dry biomass has a 50% carbon content (Goetz et al., 2009). The major focus is above-ground biomass (AGB), while below-ground biomass is usually modelled from AGB with conversion factors. AGB of a tree can be estimated through

destructive or non-destructive approaches (Köhl et al., 2006). In the destructive approach, the tree is harvested, cut down, dried, and weighed. This direct assessment of biomass produces an accurate measurement of AGB but is impractical for extensive forest inventories as it is often time-consuming and expensive. Therefore, the non-destructive approach is frequently used, in which the tree AGB is estimated using a selected empirical allometric model of some easier-to-measure variables (Brown, 1997). An allometric model is built through a destructive sampling of sample trees. These sample trees should be selected in a way that they are able to represent the population of interest in terms of *DBH* ranges, tree species, and soil conditions (Köhl et al., 2006). Many allometric models have been developed for different applications, forest types, and site conditions. The standard predictor variable is the tree diameter at breast height (*DBH*) or a combination of the tree variables, such as *DBH* and tree height (*H*) or *DBH*, *H*, and wood density (*WD*). AGB estimation using an allometric model including only *DBH* often captures most of the variability (Albert et al., 2014). However, incorporating additional predictor variables may increase the predictive performance of the model.

The most accurate method for estimating forest AGB over large areas is through sample-based forest inventory (Zhang et al., 2012). Based on the field inventory data and the selected allometric model, the estimated AGB at plot level is extrapolated to larger areas (Chave et al., 2005). However, collection of field data requires much effort, particularly in remote and highly inaccessible forests. To combat these difficulties, remote sensing data are found to be useful to retrieve information with less effort and time, repeatedly. But then, AGB estimation from remote sensing alone is hardly feasible. Field data are still needed to establish or to calibrate the corresponding AGB estimation model. The integration of field inventory and remote sensing data has been proven to be suitable to make a wall-to-wall AGB estimate over large areas in a more efficient and realistic manner (Tomppo et al., 2008). In this approach, the remote sensing response signals are combined with the field-based AGB estimates from the field inventory plots to develop an AGB estimation model.

Other studies use basal area to estimate biomass (Suganuma et al., 2006; Torres and Lovett, 2012). Basal area is a good predictor of biomass since both variables are proportional to trunk diameter. The study of Suganuma et al. (2006) found basal area, rather than the canopy cover or leaf area index, to be the best predictor for estimating biomass in open dry forests. In addition, basal area can be rapidly estimated using Bitterlich sampling, also known as angle count sampling (ACS). While implementation of ACS is often very simple, it requires a clear visibility of the trees, which cannot always be attained, particularly in a closed and complex forest structures.

1.5. Remote sensing for estimating forest biomass

A long history exists on the use of remote sensing to gather information for forest assessments, such as the spatial extent of forest cover, forest type, and biophysical and biochemical conditions at a particular site (Boyd and Danson, 2005). In line with the increased interests in quantifying forest carbon and biomass, remote sensing data across various platforms (airborne and spaceborne), sensors (optical, radar, and LiDAR), and sources of electromagnetic energy (passive, active) have been studied to enhance the estimation of forest AGB derived from field observations. Each sensor type and platform has its own characteristics and limitations.

Passive optical sensors use the electromagnetic energy emitted from the sun to provide information on vegetation. The sensors usually cover parts of the electromagnetic spectrum from visible to infrared (400-2500 nm). The information recorded by the sensor is the response signal from the interaction between the incident energy and vegetation canopy surfaces. These spectral responses depend on the sensitivity of the spectrum to the variation in the canopy structures associated with leaf structures, pigmentations, and moisture contents (Jones and Vaughan, 2010). Passive optical sensors include multispectral and hyperspectral instruments. A multispectral sensor usually has 4 to 8 bands, while a hyperspectral sensor consists of many narrow bands, which provide a continuous spectral response signal dataset (Fassnacht et al., 2016). The availability of hyperspectral data is still limited. The data is mainly available based on airborne platforms, thereby capturing small areas at high costs. On the other hand, multispectral optical data is available at various spatial resolutions from less than 1 m to more than 100 m, some of which can be acquired at relatively little or no cost.

Optical remote sensing data quality depends on the received influences from atmospheric conditions at the time of acquisition, which poses a major problem, especially for its implementation in the tropics. In most cases, to get a clear optical image without clouds and haze is very challenging since there are locations where clouds are persistent throughout the entire year. The problem can be reduced with a higher temporal resolution system that allows acquiring images at a higher frequency.

The use of multispectral data for developing AGB estimation models has been researched using different satellite data: IKONOS (Proisy et al., 2007), Quickbird (Fuchs et al., 2009), Worldview-2 (Mutanga et al., 2012), RapidEye (Pfeifer et al., 2016; Rana et al., 2013; Wallner et al., 2015), Sentinel-2 (Sibanda et al., 2015), Landsat (Lu, 2005; Lu et al., 2004; Wijaya et

al., 2010b), and MODIS (Anaya et al., 2009; Baccini et al., 2008; Gallaun et al., 2010). The studies report that this optical data can be used to provide information on the variability of horizontal vegetation structure but is limited for vertical structure. Since vertical structure information is important for biomass, relationship between the derived remote sensing variables and the field AGB estimates is limited in forests with high variability of stand structures. There are also studies that reported promising results on the use optical data of relatively high spatial resolution of ALOS AVNIR-2, WorldView-2, and IKONOS to estimate AGB, with a model fit (Adj R^2) of 0.8 for forest in southeast coast of China, (Sarker and Nichol, 2011), 0.8 for forest in northeast Madagascar (Eckert, 2012), and 0.7 for oil palm plantation in Africa (Thenkabail et al., 2004), respectively.

Besides its limitations, in comparison to other sensors, optical systems hold some advantages for use in forest AGB monitoring systems. The data are available for longer periods, at regular time intervals, at global and local scales, and in various spatial resolutions. Some of these data are available at a relatively low cost (e.g., RapidEye) or free of charge (e.g., Landsat and Sentinel-2). Advanced knowledge about data processing and analysis has been extended. Data processing tools are also available in commercial and in open source platforms.

This study focuses on data derived from passive optical remote sensing. However, it is worth mentioning data from active remote sensors that generate their own energy. For example, radio detection and ranging (Radar) or synthetic aperture radar (SAR) and light detection and ranging (LiDAR) that are also used for estimating AGB. Due to its wavelength characteristics, radar data is not influenced by clouds and haze. With the ability to penetrate through the forest canopy, radar data is expected to capture information on vertical vegetation structures. However, radar backscatter also saturates in forests with complex structures and high biomass density (Zolkos et al., 2013). It is also difficult to distinguish between vegetation types and speckles in radar data (Lu et al., 2014), which presents a major issue for forest AGB estimation. In comparison to optical data, the availability of radar data is still limited and costly (Joshi et al., 2015). Compared to optical and radar data, LiDAR has been found to be a promising tool for an accurate estimate of AGB with much higher saturation values (Næsset et al., 2016, 2015; Urbazaev et al., 2018; Zolkos et al., 2013). LiDAR has the ability to sample the vertical distribution of canopy and ground surfaces, providing detailed structural information on the target vegetation, whereas more accurate estimations of basal area, crown size, tree height and stem volume can be expected (Lu et al., 2014). However, LiDAR data usually covers only small areas and application of LiDAR for wall-to-wall AGB estimation over large areas is still cost-prohibitive (Pfeifer et al., 2016; Saarela et al., 2015).

1.6. Background of the study

1.6.1. The DFG Project (KL/894/17) of Kalimantan

This study was part of the research project "Development of an integrated forest carbon monitoring system with field sampling and remote sensing" funded by the Deutsche Forschungsgemeinschaft (DFG) under the Project number KL894/17. In this project, the Chair of Forest Inventory and Remote Sensing, Georg-August-Universität Göttingen, was collaborating with the GeoBio Center, Ludwigs-Maximilians-Universität München, Germany, and the Center for International Cooperation in Sustainable Management of Tropical Peatland (CIMTROP) at the University of Palangka Raya, Indonesia. The project aimed to contribute to a methodological improvement of forest carbon monitoring systems in tropical peat forests by integrating sample-based forest inventory and remote sensing data to develop biomass estimation model.

The DFG project took place on a logged-over secondary forest of lowland tropical peat-swamp forests in the Province of Central Kalimantan, Indonesia. In this region, a large part of the forests had been under logging concessions intensively starting from 1970s (Sorensen, 1993). In the late 1990s, about 1 million ha of forests in this region was designated to be converted into farmland under the Mega Rice Project (MRP). The project was initiated by the Indonesian government (Hirano et al., 2012; Page et al., 2002) and aimed to achieve national self-sufficiency in rice production. To support the program, large parts of the peat-swamp forests were cleared, and canals up to 30 m width with an approximate length of 4000-4500 km were constructed for the irrigation system (Jaenicke et al., 2008; Putra et al., 2008). The MRP project failed and terminated in 1999. Among the reasons of the project failure were the knowledge gaps on peatland ecosystem function, the hydrology of peatland, peat subsidence, and the effect of long term drainage on peatland (Osaki et al., 2016). The MRP project left the remaining forests and a large area of devastated peatland (Hirano et al., 2012) in a fire-vulnerable state (Page et al., 2002).

Knowing that the information on the AGB of the remaining peat-swamp forests was essential, the DFG project was conducted in this site. This information can be used to estimate the amount of potential carbon that could be emitted into the atmosphere due to such burning. This information will also support the commitment of the government of Indonesia to reduce their CO_2 emissions from forests and peatlands by year 2020 and year 2030.

1.6.2. Challenges

Field observation using angle count plot method

Fixed-area plots and angle count plots are the frequently used plot designs in forest inventories (Piqué et al., 2011; Scott, 1990). Plot designs define the rules of the inclusion of trees at a sample point. A sample plot with a complete measurement of trees is considered less efficient than that of nested sub-plots. In nested sub-plots, trees are sampled with unequal inclusion probabilities, depending on their dimension, e.g. larger trees have a higher probability to be sampled in the larger sub-plot than smaller trees in the smaller sub-plot. The idea of nested sub-plots was further developed by Bitterlich and was introduced as angle count sampling in the year 1948 (Bitterlich, 1984). The name "angle count sampling" comes from a fixed opening angle used in the device to define the sample trees. It is also known as point sampling, variable plot sampling, relascope sampling, or Bitterlich sampling. In a strict sense, it is more of a plot design method than a sampling design method and should therefore more correctly be called "angle count plot". Hence, the term "angle count plot" (ACP) is used in this thesis, along with fixed-area plot (FAP).

The ACP mechanism is simple: with a device that produces a defined opening angle, a 360-degree sweep centered at the sample point is performed (Bitterlich, 1984). All trees that appear, at 1.3 m height, larger than this angle are counted as sample trees. The inclusion probability of a tree is then strictly proportional to the tree basal area (Grosenbaugh, 1958). Larger trees are included up to a larger distance, while smaller trees are only counted at closer distances. The opening angle defines the basal area factor (BAF), with which the number of counted trees is simply multiplied to produce a per-plot basal area estimate per hectare. If the opening angle matches the tree *DBH*, this tree is a border tree. To ascertain whether the tree should be counted or not, the *DBH* of the tree and its distance to the plot center needs to be measured. As this measurement is time-consuming, a standard approach is to count every second border tree or to count a border tree half.

Due to the ease of implementation and considering that the inclusion probability of the tree is directly related to stand basal area, ACP is found to be an efficient method and is widely used in forest inventories (Piqué et al., 2011; Schreuder et al., 1992; Sukwong et al., 1971). Basal area, the sum of the cross-sectional area measured at breast height (commonly at 1.3 m) of all trees in a stand, is an important variable since it is closely related to stand volume (Hasenauer and Eastaugh, 2012; Scrinzi et al., 2015), biomass (Pérez-Cruzado et al., 2011), and carbon stock (Torres and Lovett, 2012). Earlier studies also used ACP to derive diversity indices (Motz et al., 2010; Sterba, 2008) or, in combination with remote sensing data, to feed

regionalized models of biomass (Englhart et al., 2012b, 2011; Jochem et al., 2011; Kronseder et al., 2012; Scrinzi et al., 2015) and timber volume (Scrinzi et al., 2015), for example.

Among the relevant sources of measurement errors in ACP is limited stand visibility (Bitterlich, 1984; Ritter et al., 2013), especially in forests with dense undergrowth. Trees at relatively large distances are then hardly visible. Tree omission is a typical plot measurement error that also occurs in FAP. This issue is particularly relevant for ACP since it is usually concerned with larger trees that are further away from the plot center. Selection of a higher BAF, which correspond to a "smaller plot size", can minimize this limited visibility problem. On the other hand, a higher BAF will typically lead to higher standard errors on the estimation (Becker and Nichols, 2011; Brooks, 2006).

The selection of a suitable BAF depends on the visibility conditions. However, the distance of visibility of the forest is usually unknown. Selecting a suitable BAF also constitutes a typical optimization issue of balancing precision, the probability of measurement errors, and cost. Some general recommendations of a proper BAF are given by Bitterlich (1984): a BAF of 1 or 2 is commonly used in temperate and boreal forests with good visibility, and a factor of up to 9 is common in dense tropical rainforests with limited visibility.

While this method is convenient and efficient for field data collection, its implementation (e.g., selection of a BAF) very much depends on the forest conditions. A number of studies have looked at the identification of a suitable BAF under different forest types (Becker and Nichols, 2011; Bitterlich, 1984; Brooks and McGill, 2004). But to the author's knowledge, none of the studies have been conducted specifically on tropical peat-swamp forests or by explicitly using the distance of visibility in the forest as basic information in selecting a suitable BAF. Tropical peat-swamp forest is characterized with dense undergrowth vegetation, in which obtaining error-free observations from ACP poses an issue. Some studies that employed ACP in this forest type did not specify the BAF used for their study or only reported that it was "relascope sampling with a wide scale of BAF" (Englhart et al., 2012b, 2011; Kronseder et al., 2012). Therefore, this study was conducted to examine some of the remaining questions. For example: How to define the maximum distance of visibility for implementing ACP? What is the effect of restricted visibility in the estimation based on ACP? What is the suitable BAF for implementing ACP in a closed and dense forest and in an open forest?

Combining field sample plots and optical remote sensing data

Remote sensing cannot directly measure forest AGB. Auxiliary information is usually extracted from data at an area where ground reference data is available, e.g., forest inventory sample plots. This auxiliary information is then used as potential predictor variables of AGB. They linked with the AGB estimates of the plot, based on field observation, to develop an AGB estimation model (Næsset et al., 2016). The AGB of the area of interest is estimated using the developed model. The combination of the two datasets allows mapping the AGB distribution of the entire area of interest since field sample plots alone do not describe the spatial distribution.

To date, linking remote sensing predictor variables and field-observed AGB presents many challenges. Among the potential predictor variables of AGB derived from optical remote sensing data are the spectral signatures, vegetation indices, and texture features (Lu et al., 2014; Sarker and Nichol, 2011).

Much research on the use of optical remote sensing for AGB estimation has been conducted. This research reported various results on the relationships between remote sensing predictor variables and field-observed AGB, depending on the data and the selected approach, as well as the relevant forest characteristics. While a number of comparative studies on the remote sensing AGB estimation model using different sensors and approaches have been done, there are still limited studies comparing the relationships between the predictors and AGB in different forest characteristics, particularly in tropical peat-swamp forests. Some primary reasons for this limited research for this specific forest type are related to the difficult accessibility of the forest, restricted availability of cloud-free remote sensing data, as well as the complex stand structures and high diversity of tree species compositions. Therefore, research in comparing the relationships between the predictor variables from optical remote sensing data and the field-observed AGB of tropical peat-swamp forests and other forest types is needed.

This study developed AGB models between field-observed AGB and RapidEye satellite data in a tropical peat-swamp forest and a temperate forest. The high spatial resolution of RapidEye data is expected to capture more of the AGB variabilities in the sample plots, which in turn may increase the relationships between the two variables. The relatively high temporal resolution of 5.5 days of RapidEye data is advantageous to reduce the issue of cloud cover in the tropics. The sensor also covers the portion of the red-edge band that is found to be more sensitive to vegetation conditions (Barnes et al., 2000; Gitelson and Merzlyak, 1994, 1997). It is expected that the vegetation indices developed by using the combination of red-edge and the

GLCM texture features derived from red-edge, will show higher relationships with AGB in both forests. Studies using RapidEye data to develop such an AGB estimation model have been conducted with promising results in tropical forests of Borneo Island (Englhart et al., 2012a; Pfeifer et al., 2016), also in Miombo Woodlands, Tanzania.

In the study of Næsset et al. (2016), model-assisted estimators were used in the estimations. Such estimators allow making inferences on the developed model. The requirement of using such estimators is the availability of the field sample plots based on the probability sampling approach, in which the model prediction is corrected using the deviation between the model prediction and the field reference (Saarela et al., 2015). By using the model-assisted approach, the efficiency of combining the remote sensing and the reference field data in the AGB estimation, relative to the estimation from field data alone, can be calculated. In the comparative study of Næsset et al. (2016), they reported a higher relative efficiency value from RapidEye of 3.3, than that of the (InSAR) TanDEM-X data. The study was conducted in the Miombo Woodland of Liwale, Tanzania, using cloud-free RapidEye imageries. More research is needed to study the efficiency from RapidEye-assisted estimates of mean AGB estimates in other forest types.

Given these key advantages, this study was conducted to assess the applicability of RapidEye data for AGB estimations in two different forest types using the model-assisted estimators.

Field plot size for AGB estimation using optical remote sensing data

The estimated AGB of the area of interest depends on the ability of the model to relate the AGB estimates of the plot with the auxiliary information on the entire area of interest. In this regard, a perfect co-registration between the field sample plots and the remote sensing data is required to ensure accurate estimators of the target estimation (Saarela et al., 2016). This co-registration effect is reported to be less in larger plots (Næsset et al., 2015). Comparing the advantages between small versus large plots, the use of larger plots increases the chance of getting more overlapped area between the two data (Gobakken and Næsset, 2008). As a plot gets larger, the field and the remote sensing data also cover more of the spatial variations, in which a decreasing model error is expected (Zolkos et al., 2013). A larger plot is also likely to reduce the effect from the boundary-related elements, as the ratio between perimeter to area is decreased. The ratio between perimeters to a given plot area also depends on the geometric shape of the plot.

When remote sensing data is to be included in the estimation of field data, the selection of the field plot size should also consider the variability of the variable of interest to be captured by the remotely sensed data. It is found to be necessary for the forest inventories design using remote sensing auxiliary dataset (Næsset et al., 2015; Tomppo et al., 2017). However, field data collection is costly. Finding an optimum plot size is always an issue, as well as determining an efficient method for fieldwork. There is always a trade-off in selecting a plot size; smaller plots with higher sample size, or larger plots with lower sample size to reach a certain precision of the target variable of interest. Larger plots tend to capture more of the variability within the plots and reduce the variance among plots. On the other hand, smaller plots tend to capture more of the variability across plots, which contribute to higher uncertainty of the estimation, depending also on the sample size.

Despite its importance, there is not much research on the effect of different plot sizes and shapes in AGB model performances. There were some studies using LiDAR (Mauya et al., 2015; Zolkos et al., 2013) and radar (Saatchi et al., 2011), but to the author's knowledge, not for optical remote sensing data. This study was conducted to analyze the effect of different plot sizes and shapes on the AGB model performance derived from RapidEye data. As larger plot sizes always mean higher costs, the analyses of the gain that can be achieved by the inclusion of remote sensing data by enlarging the field observation plots needed to be calculated. This study calculated the gain in precision of the AGB estimate from the inclusion of RapidEye data and compared to the precision of the estimate derived from using the field data only.

1.7. Objectives of the study

1.7.1. Overall objective

The overall objective of this study is to contribute to the methodological basis of the integration of field inventory and optical remote sensing data of high spatial resolution for forest AGB estimation.

1.7.2. Specific objectives and research questions

Based on the overall objective, the study focuses on three relevant research topics as sub-studies (Fig. 1.2); (i) research into the effect of forest restricted visibility in the basal area estimates from Angle Count Plot (ACP); (ii) research into the integration of field inventory

and optical RapidEye data for AGB estimation model; and (iii) research into the effect of field plot sizes on the AGB estimation model derived from field inventory and RapidEye data. Each of the sub-studies aims to contribute to the methodological development of forest AGB estimation, whereas in the sub-studies (ii) and (iii), the AGB estimations were derived by combining field and remote sensing data.

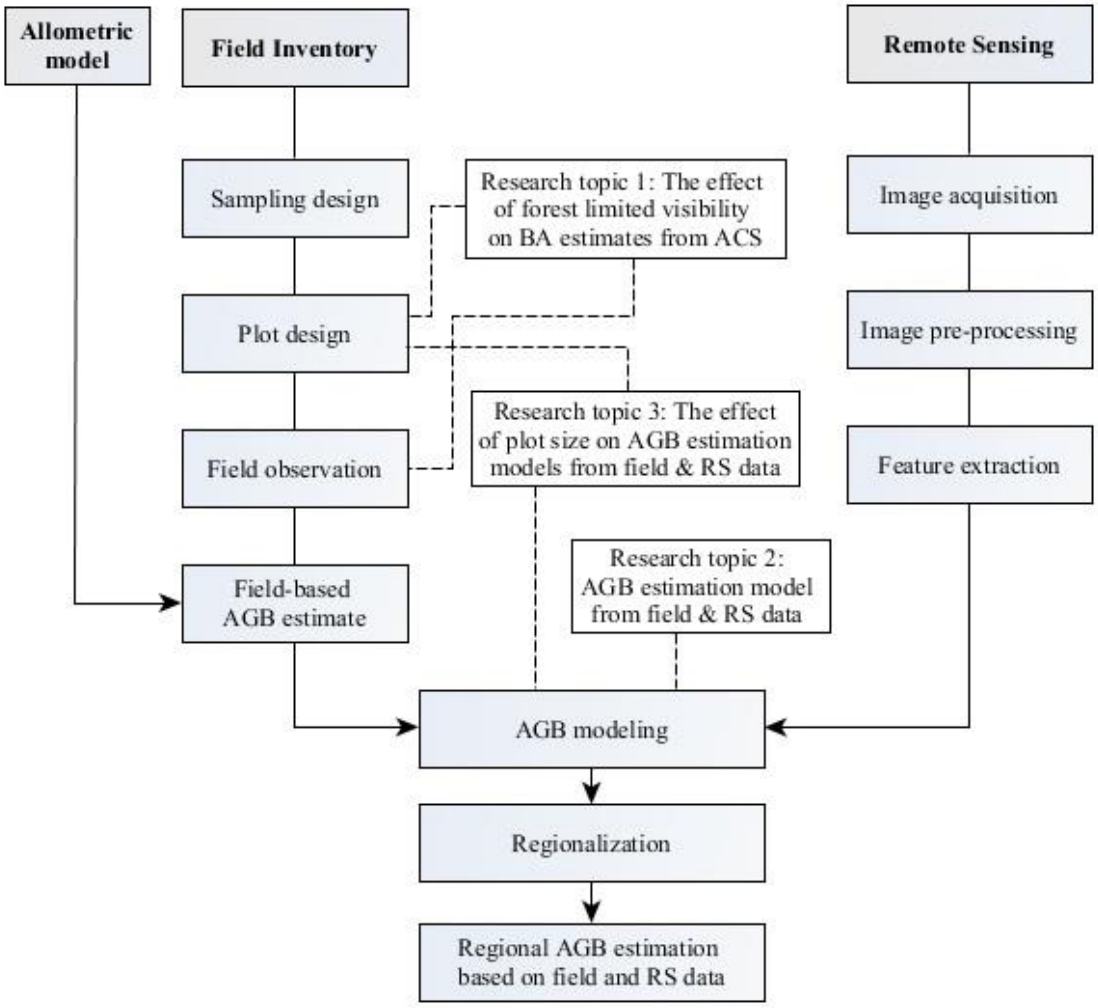


Fig. 1.2. Contribution of the study to forest above-ground biomass (AGB) estimation using the field observation and the remote sensing data. BA = basal area, ACP = angle count plot, and RS = remote sensing.

Each sub-study deals with a specific objective and the following research questions.

1. Determining a distance of visibility and a suitable basal area factor to implement ACP.

This sub-study contributes to the determination of basal area factor (BAF) based upon the given forest conditions when the ACP approach is employed for field data collection. Data

from two different forest types, a dense tropical peat-swamp forest of Sebangau and an open savanna forest of Ncaute, were used and compared for the study.

To achieve the goal, some research questions were formulated:

- How to define the distance of visibility for ACP?
- How to determine a suitable BAF for ACP in a forest with a visibility issue?
- What are the consequences for selecting a higher/lower BAF than the suitable one?
- What is the effect of forest restricted visibility on basal area estimates from ACP?

2: Developing and comparing the AGB estimation models using field inventory and optical RapidEye data in two different forest types.

This sub-study was conducted to improve the understanding of the relationship between field AGB estimates and RapidEye derived predictors. By using the same approach, this study compares the results from two different forest types, a tropical peat-swamp forest of Sebangau, Indonesia, and a temperate forest of Hainich, Germany. Further comparison of the relationships between field AGB estimates and RapidEye predictors were based on two different field datasets, each with a different plot design, in the Hainich study site.

For this sub-study, the following research questions will be addressed:

- How are the relationships between field AGB estimates and predictor variables derived from RapidEye data in each dataset?
- What is the AGB estimation model derived from each dataset, and how is the performance of each model?
- What is the efficiency of the inclusion of RapidEye data in AGB estimation relative to the estimation from field data only, across the datasets?

3: Determining the effect of different field plot sizes on the AGB models from field inventory and RapidEye data.

This sub-study was conducted to improve the understanding on the effect of field plot sizes to the performance of AGB models derived from the field and the RapidEye data. The dataset of a large full census plot from a temperate forest was used to create different datasets of different plot sizes.

For this specific objective, the following research questions are addressed:

- What is the effect of different plot sizes on the field AGB estimates and the corresponding standard error?
- What is the effect of different plot sizes and shapes on the performance of AGB estimation models derived from the field and RapidEye data?
- What is the efficiency of the inclusion of RapidEye data in AGB estimation, relative to the estimation from field data only?

Chapter 2

Materials

2.1. Study area and field inventory in the Sebangau

2.1.1. Description of study site

The study was conducted in part of the Natural Laboratory of Peat-Swamp Forest (NLPF) of Sebangau, geographically located between 113° 46' to 114° 02' E and 2° 16' to 2° 31' S, Province of Central Kalimantan, Indonesia (Fig. 2.1).

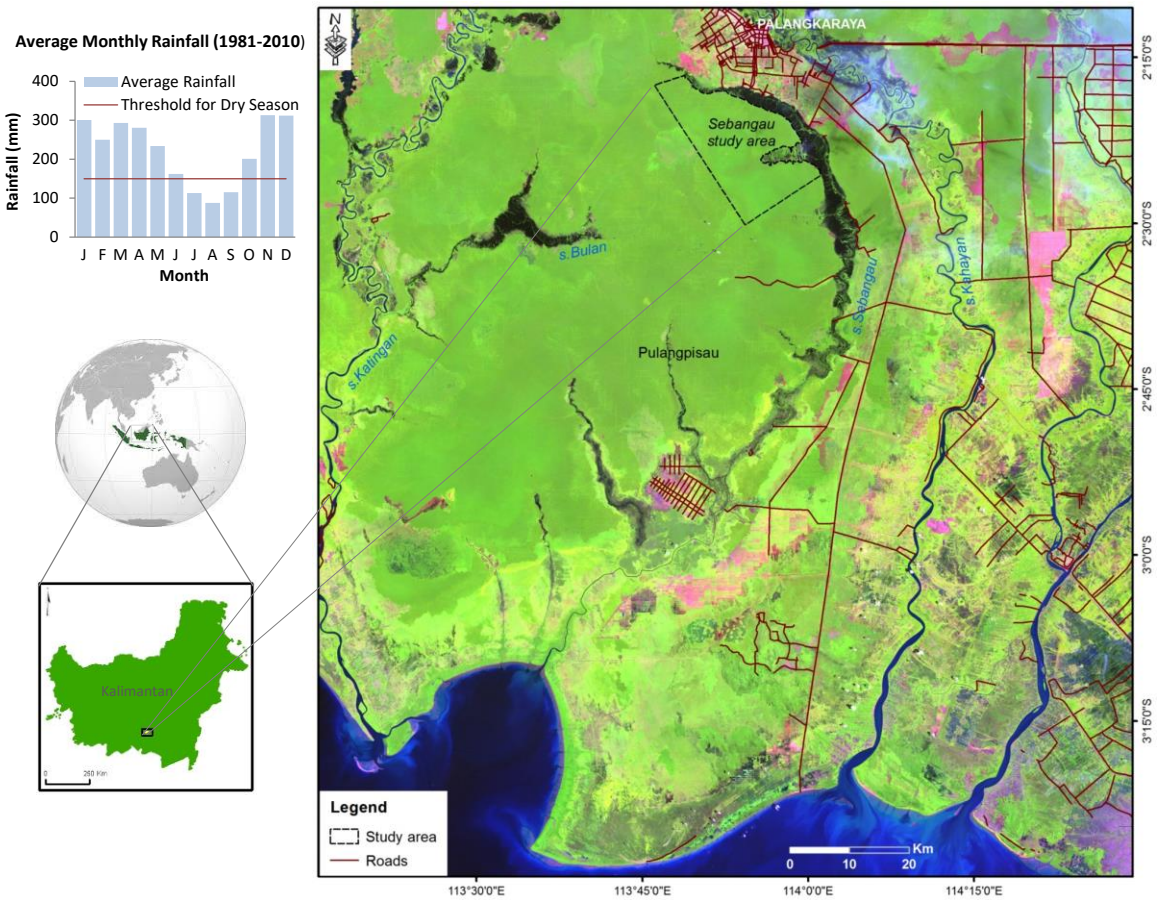


Fig. 2.1. The study area of Sebangau peat-swamp forest and surrounding lands. Background image is Landsat 5 imagery, acquired on 10th February 2010, presented in band combination of short-wave infrared, NIR, and red in the red-green-blue color composite (RGB 543). In addition, information on mean monthly rainfall (1981-2010) is given.

The lowland peat-swamp forests in the study area are situated on flat plains with an elevation of 10-30 m a.m.s.l. (Page et al., 1999), between the Sebangau and Katingan rivers of the southeastern part of Borneo Island (Fig. 2.1). In this region, the climate is characterized by tropical wet weather (Peel et al., 2007), with a mean annual rainfall (1981-2010) of 2600 mm (BMKG, 2014) and a mean annual air temperature (2002-2010) of 26.2°C (Hirano et al., 2014). Most parts of the study area are peat-swamp forest with a high diversity of tree species (Page et al., 1999). The peat has taken more than 10,000 years to accumulate mass (Page et al., 1999), with peat layers measuring up to 10 m (Jaenicke et al., 2008; Page et al., 1999). The forest was heavily logged for 25 years (Campbell, 2013) before it was transformed into the NLPF in 1997 for scientific research purposes under the management of the Center for International Co-operation in Management of Tropical Peatland (CIMTROP) of Palangka Raya University (Morrogh-Bernard, 2009). Fig. 2.2 shows the forest conditions during the fieldwork in 2013-2014.

Fig. 2.2. (a) The abandoned railway previously constructed by concession loggers to access the forest. (b) The forest floor is waterlogged, seasonally flooded, or flooded along the year. The water is dark brown, indicating a high content of organic matter. (c) Trees with stilted roots as adaptation to wet environments. (d) Inside the Sebangau forest.

Given the unique ecosystem characteristics, including the diversity of flora and fauna, the forest became part of Sebangau National Park by a government decree in 2004 (Indonesia MoF, 2004). However, there is still illegal logging in parts of the forest, according to oral reports from the forest guards of CIMTROP.

2.1.2. Sampling and plot designs in the Sebangau

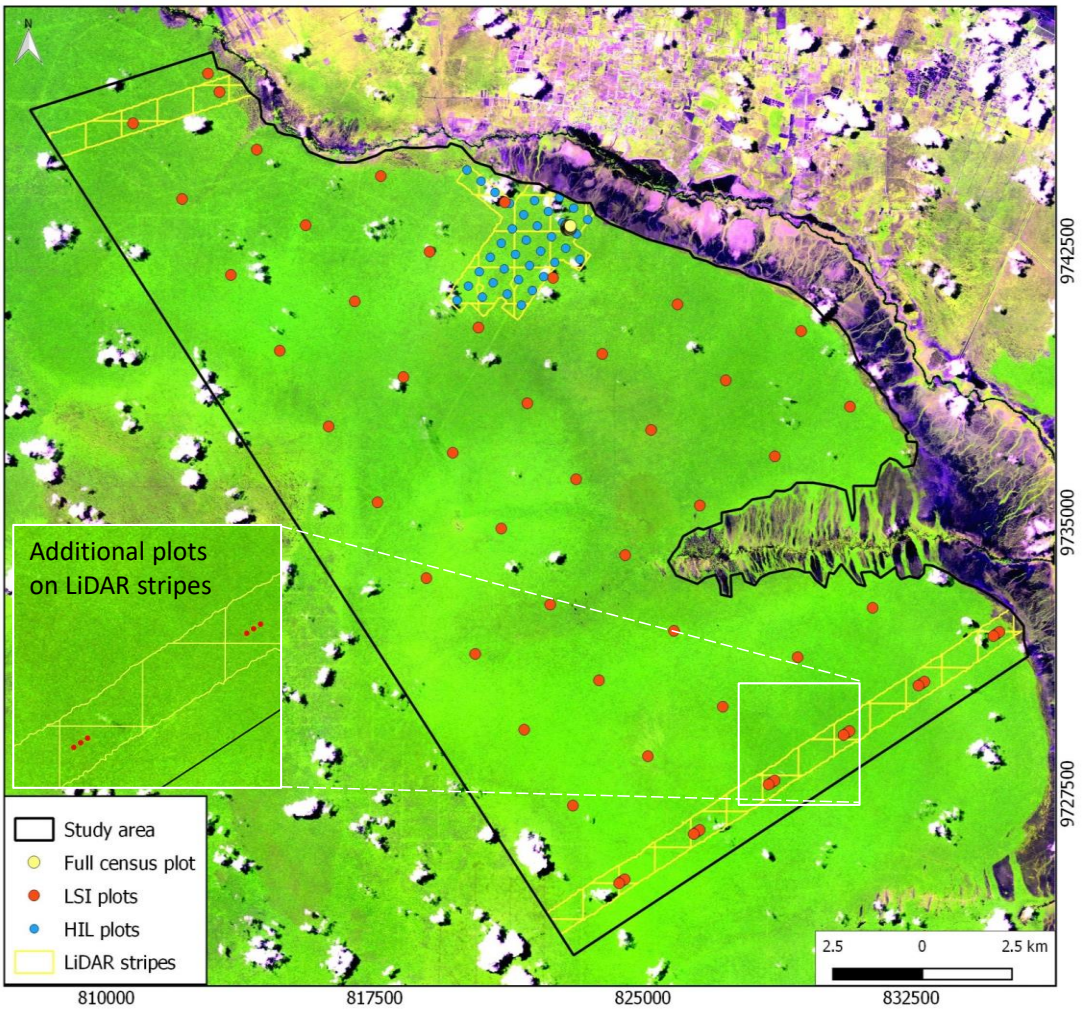


Fig. 2.3. Three designs of data collection on the Sebangau study area; full census (FC) plot in yellow dots, large scale inventory (LSI) plots in red dots, and high intensity LiDAR (HIL) in blue dots. Two additional plots were added for each LSI plots located on the LiDAR sample (index). Background image is RapidEye imagery, acquired on 20th of September 2013, presented in band combination of red-edge, NIR, and red in the red-green-blue color composite (RGB 453).

Fieldwork in Sebangau was carried out from July 2013 until June 2014. The field conditions were extremely challenging as this particular area of the forest is very difficult to access. Almost none of the remaining rail tracks from the era of the concession company were

functional anymore (Fig. 2.2). The interior of the forest was only accessible by foot with an average walking speed of 2.5 km per day as it was experienced in this study.

Three different sampling approaches guided by different project objectives were implemented independently to collect field data (Fig. 2.3). The three approaches were: (i) Large Scale Inventory (LSI) design over the entire study area (31,113 ha), (ii) High Intensity LiDAR (HIL) design over the LiDAR data covered area (869 ha). This HIL frame encompassed the largest patch of LiDAR data. It did not consider the other two LiDAR stripes of 5 and 15 km lengths located 25 km from each other. Notice that the LiDAR data were not part of these analyses, but rather the in-field data acquired within the LiDAR-based area frame were used. (iii) a Full Census (FC) design over one large observation plot (1.44 ha). In this one large plot, full enumeration was constructed to allow for simulation studies. Both the HIL and FC designs fell within the LSI frame (Fig. 2.3).

The LSI and HIL plots were sampled probabilistically with systematic grids of 2500 m and 500 m, respectively. The systematic design was selected for the uniform coverage of samples over the study area and for the generally higher effectiveness in comparison to the simple random sampling design with the same number of observation plots (Köhl et al., 2006; Van Laar and Akça, 2007). The grid sizes were determined considering both target precision and available resources. The resulting sample sizes of LSI and HIL for these grid sizes were $n = 46$ and $n = 35$, respectively. The azimuth angle of the LSI grid was 57°, as defined by a random number. On the other hand, the HIL grid followed the orientation of the LiDAR polygon. In addition, $n = 13$ plots were taken to cover more observations on the LiDAR stripes. For each of the six LSI plots located on the southern LiDAR stripe, two additional plots were added. These 12 additional plots were located at 100 m distance from each side of the original plot and aligned with the LiDAR stripe (Fig. 2.3). Another plot was located on the northern LiDAR stripe, about 600 m below the LSI plot. The total number of LSI plots was $n = 59$ (Table 2.1).

Table 2.1. Three sampling designs in Sebangau: Large Scale Inventory (LSI), High Intensity LiDAR (HIL), and Full Census (FC). In the LSI design, additional plots ($n = 13$) were taken to augment sample size used in LiDAR analysis (see details on the text).

Variables	**LSI**		**HIL**	**FC**
	Systematic	Additional		
Area (ha)	31,113	-	869	1.4
Grid size (m)	2,500	-	500	-
Sample size (n)	46	13	35	1

The LSI and the HIL shared the same plot design: a concentric design with three circular sub-plots of radius 4, 8, and 16 m (50.3 m^2, 201.1 m^2, and 804.2 m^2), as illustrated in Fig. 2.4.

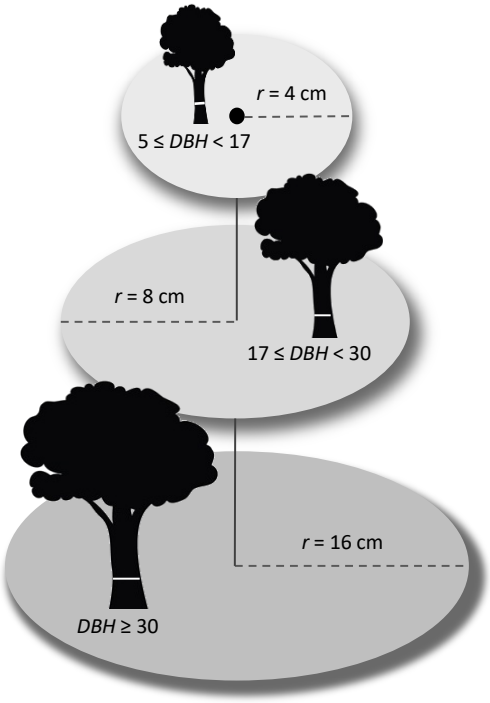


Fig. 2.4. The plot design of LSI (n = 59 plots) and HIL (n = 35 plots): concentric circular design with 3 nested sub-plots.

A circular plot is generally preferred since only one control point, the center, is required to establish the plot in the field (McRoberts et al., 2013b). The concentric design was selected to obtain efficient field measurements. Small trees were measured in smaller sub-plots, while larger trees that contribute greatly to the total AGB estimates were measured in larger sub-plots. In the smallest sub-plot, all trees with a $DBH \geq 5$ cm were selected. Trees with a $DBH \geq 17$ cm were selected in the medium sub-plot, while trees with a $DBH \geq 30$ cm were selected in the largest sub-plot. The radius of each sub-plot, and the *DBH* thresholds for each sub-plot, were determined by the project using an approach given in APPENDIX A.

The FC observation plot had a size of 120 m x 120 m, in which all trees with a $DBH \geq 5$ cm were selected. For practical reasons in the field, such as visibility and measurement efficiency, the FC plot was divided into 64 quadrants of a 15 m x 15 m that were permanently marked with plastics pipes. A reference point was established in each quadrant. While the ideal position of the reference point was at the quadrant center, some centers were located in obstructed or inundated positions. In such cases, reference points were shifted to the nearest

feasible location in the vicinity of the obstructed or inundated position, rather than being in the center of the quadrant (Fig. 2.5).

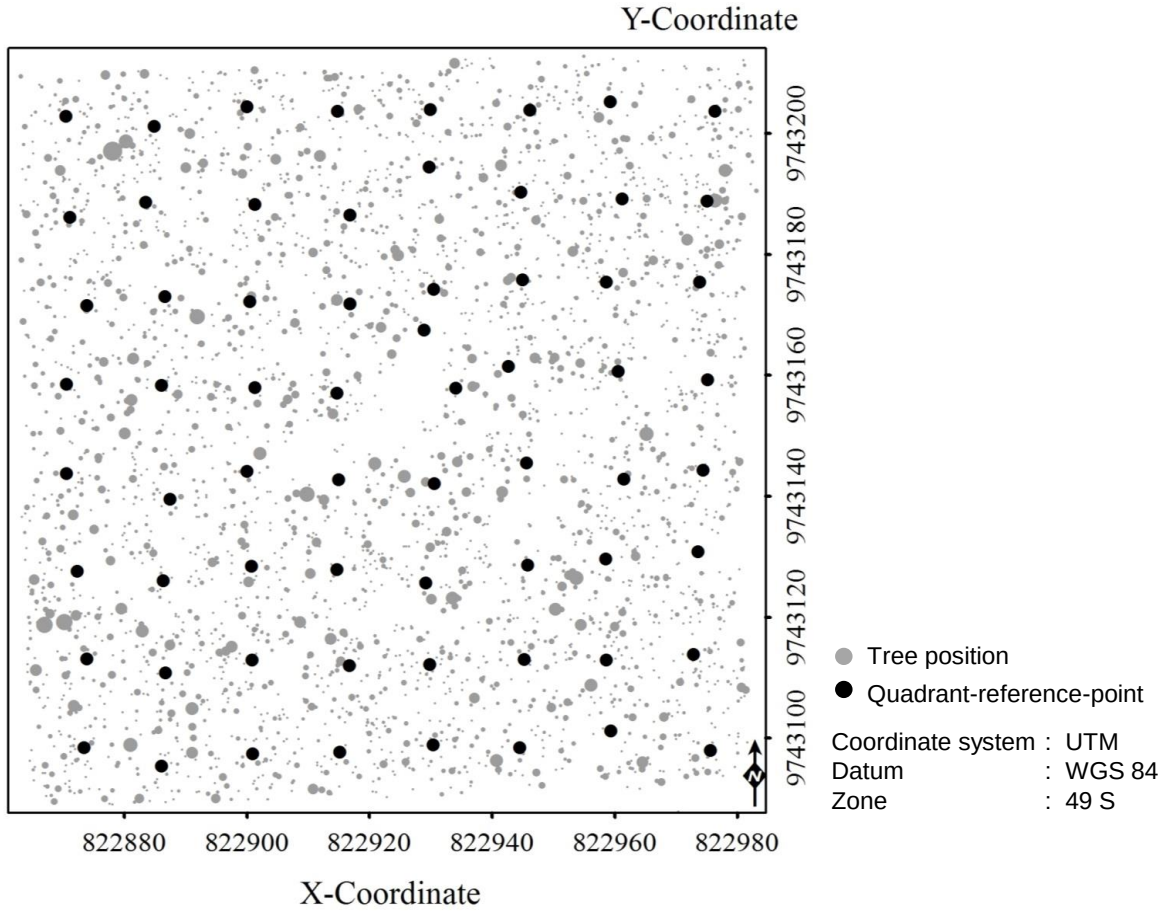


Fig. 2.5. Tree map of the FC plot, consisted of 64 quadrants of 15 m x 15 m size. The black dots are the quadrant-reference-points, from where: (i) the position of all trees in a quadrant was measured; (ii) basal area was estimated using the ACP of BAF 1, 2, and 4. The gray dots represent trees with a $DBH \geq 5$ cm (z=3427); the size of each gray dot is proportional to the tree *DBH*.

2.1.3. Field measurements

The field location of each LSI and HIL plots was found using a Garmin handheld GPS. Once the plot location was identified, final coordinates of the plots were recorded using long-term GPS measurements with an average maximum horizontal positioning error of 3 m. In each sub-plot, the following attributes were measured for each selected tree: *DBH*, tree position relative to the plot center (azimuth and horizontal distance), and local name. Measurements of height and crown-width were limited to ten trees per plot, making fieldwork more effective. These ten trees consisted of the six trees with the largest *DBH* and the four trees closest to

center of the plot. The crowns of the largest trees were expected to be more exposed and visible from above, thus easily detectable in remote sensing data. The four closest to the center were tagged to facilitate location of the plot during re-measurements. Basal area was estimated using ACP of BAF 1 from the plot center. To assess the effects of visibility on basal area estimates using ACP, some of the plots were re-assessed using BAF 2 and 4. The local tree names in the HIL and the LSI plots were recorded for the target trees that were selected according to the plot design.

On the FC plot, all selected trees ($DBH \geq 5$ cm) were tagged and measured for the same tree attributes mentioned above (*DBH*, tree position, local name, height, and crown width). The tree position was measured relative to the quadrant-reference point. ACP of BAF 1, 2, and 4 was also performed from each of these reference points. In total, $z = 3427$ stems were recorded in the FC plot as shown in the trees map of Fig. 2.5. In a separate field study, the species name of each tree in FC plot was identified by a local botanist. The results were used to identify the species name of all trees in the Sebangau study area. In case of trees in which species name could not be identified, they were classified as not available (NA) tree species.

The *DBH* of all trees in this study were measured at 1.3 m height with a diameter tape to the nearest millimeter. The definition of 1.3 m tree-height used in the fieldwork is given in APPENDIX B. The diameter of irregular stems, e.g. of a buttressed or a stilt-rooted tree, was measured above the stem irregularities as the diameter over deformation (DOD). The height where DOD was measured was also recorded. For the LSI and HIL plots, the relative tree position was recorded as the polar coordinates from the plot center. However, for the FC plot it was recorded from the quadrant-reference-point. Horizontal distance from the plot center, or from the quadrant-reference-point, to the tree was measured with a Vertex IV to the nearest centimeter. The azimuth was measured with a Suunto compass to the nearest sexagesimal degree. The tree height was measured with a Haglöf Vertex IV to the nearest decimeter. The ACP of BAF 1, 2, and 4 were carried out by using a Kramer Dendrometer (Kramer and Akça, 1995). Following the general method of ACP, border trees were counted as half.

In all designs, when a standing dead sample tree was encountered, only *DBH* and tree position were measured. General information about the fieldwork such as date and time, the plot site conditions including management level and crown closure, were also recorded.

2.2. Study area and field inventory data of Ncaute, Namibia

2.2.1. Description of study site

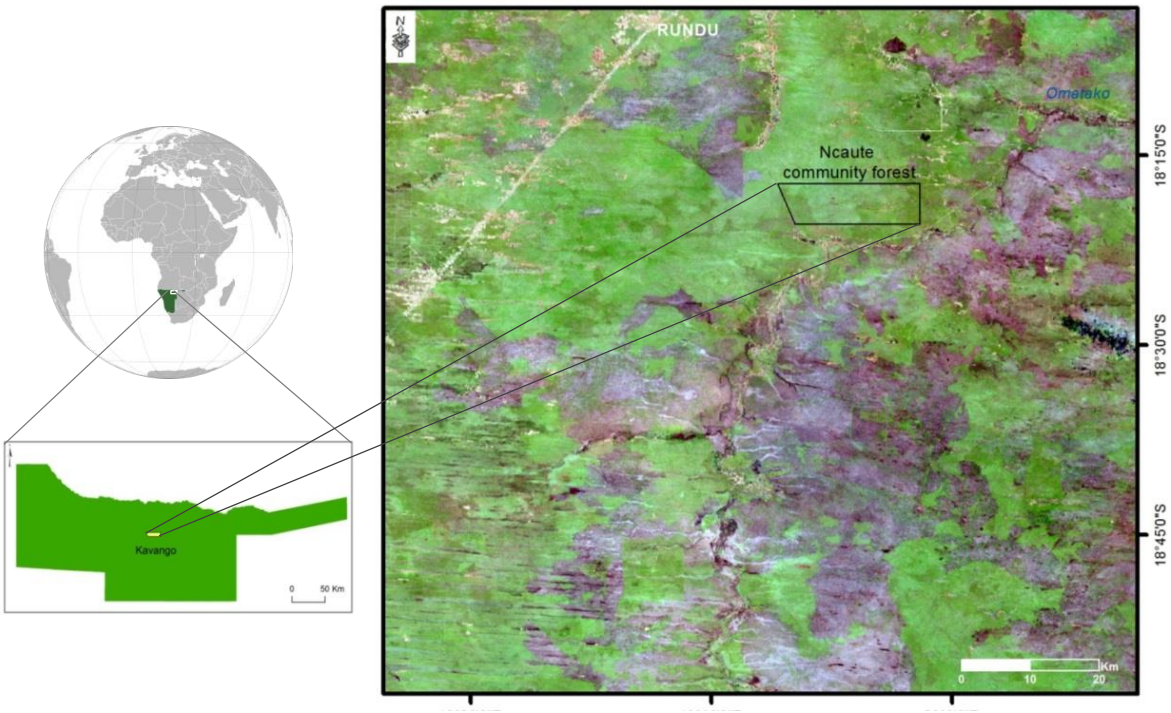


Fig. 2.6. The map of Ncaute community forest and surrounding lands. The background image is Landsat 8 imagery, acquired on 16th October 2014, presented in band combination of short-wave infrared, NIR, and red in the red-green-blue color composite (RGB 654).

The Namibian study site is part of the Ncaute Community Forest, a dry open savanna forest. This study site was chosen in compliment to the Sebangau study site for a comparative study on the effect of forest visibility in basal area estimation using ACP. The forest is located in the West Kavango region, about 30 km in the south of the regional capital of Rundu (Fig. 2.6). The region is generally flat with an average altitude of 1100 m a.m.s.l. (Erkkilä and Siiskonen, 1992). The study area is located geographically between 19° 45' to 19° 58' E and 18° 17' to 18° 21' S. The area was declared as a community forest in 2006 (Namibia MoAWF, 2006) and classified as savanna forest on sandy soils (ferralic arenosols). The climate is characterized by an annual mean temperature between 20-22°C with annual average precipitation of 500-550 mm according to (Mendelsohn et al., 2002). Fig. 2.7 shows the forest conditions during fieldwork.

Fig. 2.7. Trees in the Ncaute community forest.

2.2.2. Field measurements

The field data were collected in September 2014 by a cooperation project between the University of Göttingen, Germany, the Stellenbosch University of South Africa, and the Polytechnic University of Namibia. In total, n=33 sample plots were measured for this study. Each plot consisted of three circular sub-plots of radius 10 m, 20 m, and 30 m (314.2 m^2, 1256.6 m^2, and 2827.4 m^2). Compared to the sub-plots in Sebangau, the size of sub-plots in Ncaute study were relatively large due to the low tree density of the forest. All trees with a $DBH \geq 5$ cm were measured in the smallest sub-plot, while trees with a $DBH \geq 20$ cm and trees with a $DBH \geq 45$ cm were measured in the medium and largest sub-plots, respectively. Each selected tree was measured for the same attributes (*DBH*, tree position, and species) using the same devices and procedures as in Sebangau. As done in Sebangau, basal area was also estimated from ACP of BAF 1, 2, and 4 with the Kramer Dendrometer.

2.3. Study area and field inventory data of Hainich, Germany

2.3.1. Description of study site

The Hainich study site is located in the Hainich-Dün region of western Thuringia, Central Germany. This study site was chosen together with the Sebangau study site for the development of an AGB estimation model from field and remote sensing data. Additionally, a dataset from this study site was also used to study the effect of different plot sizes in AGB model performances. Geographically, the site is situated in between 10° 16' to 10° 33' E and 51° 01' to 51° 14' N (Fig. 2.8). The southern part of the forest belongs to the Hainich National Park as established by law in 1997 (Mölder et al., 2006), from a decommissioned military training area (Kharazipour et al., 2009). This deciduous temperate forest is dominated by beech (*Fagus sylvatica*) on mainly Luvisols and Stagnosols soil types (Fischer et al., 2010). The climate is characterized by an annual mean temperature between 6.5-7.5°C and annual average precipitation of about 750-800 mm (Birkhofer et al., 2012). Fig. 2.9 shows the Hainich forest conditions.

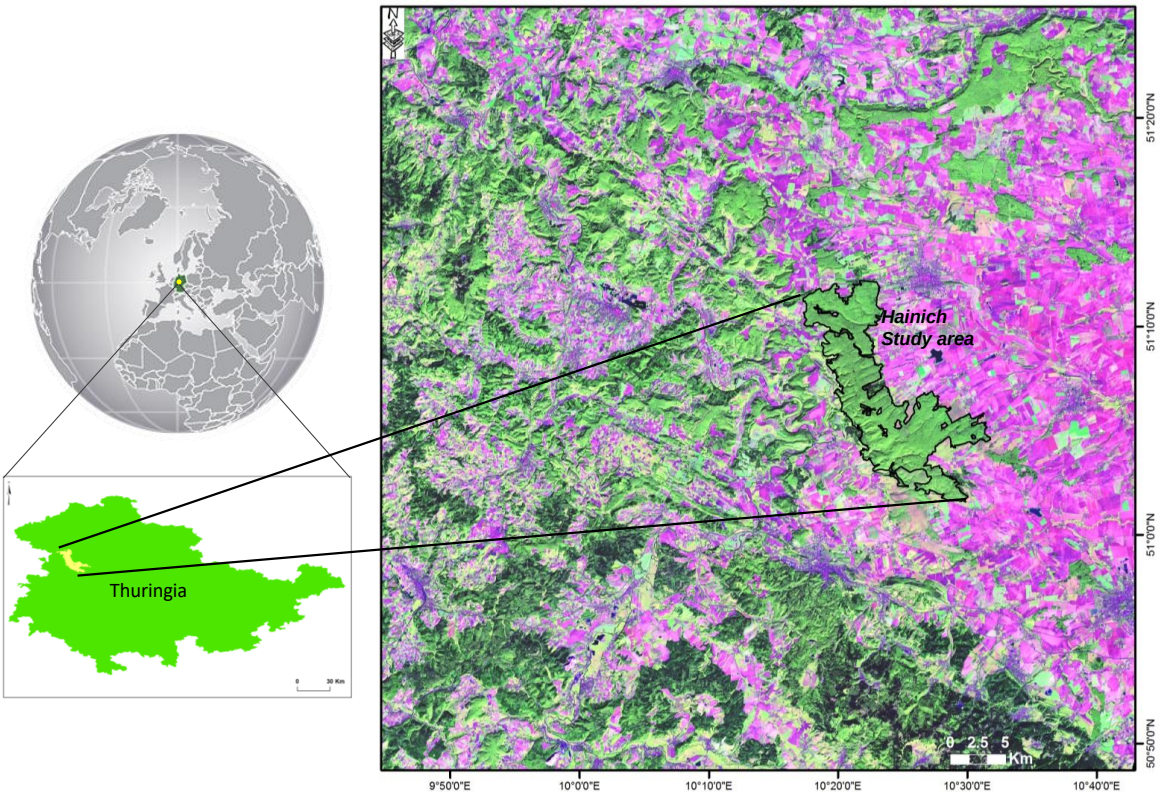


Fig. 2.8. The map of Hainich forest study area and surrounding lands. Background image is Landsat 8 imagery, acquired on 1st October 2015, presented in band combination of short-wave infrared, NIR, and red in the red-green-blue color composite (RGB 654).

Fig. 2.9. Forest canopy and forest condition in the study area of Hainich forest.

2.3.2. Field measurements

There were two field inventory datasets available for the Hainich study. One dataset referred to as experimental plots (EP), was collected under the framework of the Biodiversity Exploratories project and funded by the German Science Foundation (DFG) Priority Programme 1374 (Fischer et al., 2010). The selection of EPs was done according to Fischer et al. (2010) following a simple random sampling (SRS) design. The EPs have a size of 1 ha with a square shape of 100 m x 100 m, whereas all trees with a $DBH \geq 7$ cm were measured (Table 2.2). A total of $n = 28$ plots was collected from 2014 to 2015.

Another dataset, so called the AWF plots (AWF), was designed and measured by the AWF department of Georg-August University Göttingen (GAUG) and the Hainich National Park Administration (HNPA). AWF plots sampled the Hainich study area in a systematic grid of 800 m using different plot designs (Table 2.2). Different plot designs were used between the sample plots located inside the Hainich National Park, hereafter referred to as AWF-HNPA, and the plots located outside the national park, hereafter referred to as AWF-GAUG. To have

a comparable observation plots for the analysis, the AWF-HNPA plot design was adjusted to the plot design of AWF-GAUG. Thus, for each AWF-HNPA plot, all trees were excluded if they were located at a distance greater than 12.68 m from the plot center or they had a *DBH* < 30 cm at a distance further than 7.98 m from the plot center. Tree height was measured for approximately 5 trees per plot. The measurements were used to establish a height curve for predicting the heights of the non-height-measured trees. The AWF plots were measured during 2011-2013 and were in total n = 177 plots.

Table 2.2. Sampling and plot designs of the Hainich datasets. EP=Experimental Plot, GAUG= Georg-August University Göttingen, HNPA=Hainich National Park Administration, SRS*= Simple Random Sampling according to Fischer et al. (2010).

Plot information	**EP**	**AWF**	
		GAUG	HNPA
Sampling design	SRS*	Systematic	Systematic
Design grid size (m)	-	800	800
Plot shape	Square	Concentric circular	Concentric circular
Plot size (m^2)	10000	-	-
- sub-plot 1	-	200	200
- sub-plot 2	-	500	500
- sub-plot 3	-	-	1,000
DBH thresholds (cm)	≥ 7	-	-
- sub-plot 1	-	≥ 7	≥ 7
- sub-plot 2	-	≥ 30	≥ 25
- sub-plot 3	-	-	≥ 50
Sample size (n)	28	108	69

2.4. Remotely sensed data

2.4.1. Characteristics of RapidEye sensor

RapidEye imagery was used in two of the case studies: (i) development of an AGB estimation model from field and remote sensing data, and (ii) analysis of effects of plot sizes on AGB model performance. RapidEye is a German Earth observation satellite that was launched in August 2008 by the RapidEye AG and is currently owned and operated by Planet Labs, Inc. The RapidEye space system is a constellation of five satellites, orbiting at an altitude of 630 km. Each of the satellites is equipped with identical sensors that are calibrated equally to

each other (BlackBidge, 2015). The temporal resolution of each satellite for capturing the same location is 5.5 days. However, with a pointing capability up to ± 20º field in any direction of the observation camera, the sensor is able to make a daily acquisition of the same location. The original spatial resolution of the data is 6.5 m at nadir, which is resampled to 5 m in the orthorectified products. The data are collected in 12-bit radiometric resolution and converted to 16-bit during the on-ground processing (BlackBidge, 2015).

The RapidEye sensor is a multi-spectral push-broom imager with five spectral bands ranging from Blue (B1: 440-510 nm), Green (B2: 520-590 nm), Red (B3: 630-685 nm), Red-edge (B4: 690-730 nm), to Near-Infrared (NIR) (B5: 760-850 nm) as depicted in Fig. 2.10. RapidEye is the first commercial satellite system that included the red-edge band. While vegetation chlorophyll strongly absorbs the electromagnetic radiation (EMR) in the red band, leaf cell structures lead to strong reflectance in the NIR band. The red-edge band is located between the red and NIR bands. It is part of the spectrum where vegetation reflectance increases considerably from the red to the NIR wavelength. Additional sensitivity to detect different levels of vegetation chlorophyll content and leaf structure is expected from the red-edge reflectances (Barnes et al., 2000; Gitelson and Merzlyak, 1994, 1997). The variations shown by the red-edge region may provide additional information about the vegetation status, e.g., condition of health and nutrition (Armon and Hänninen, 2015; Eitel et al., 2011), useful for the characterization and separation of vegetation types.

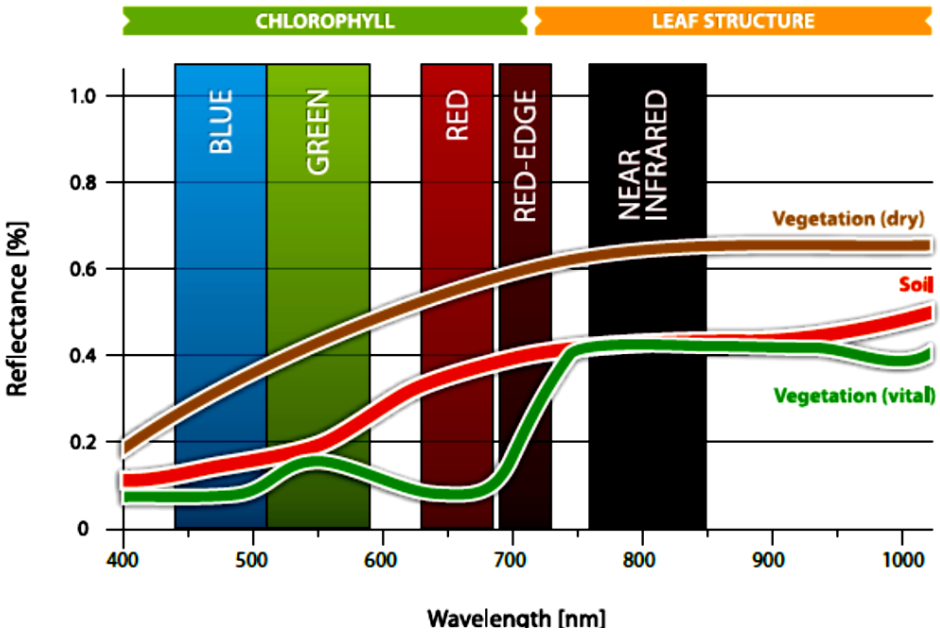


Fig. 2.10. Spectral characteristics of RapidEye images (BlackBridge, n.d.).

2.4.2. Image availability

For the Sebangau study site, a new acquisition of RapidEye data was requested from the RapidEye Science Archive (RESA) of the German National Aeronautics and Space Research Centre (DLR). The request was made for the time-series RapidEye ortho product of a processing level 3A (L3A), spanning from June 2013 to May 2014 with a time frame of 5 days. The obtained L3A data was already orthorectified and radiometrically pre-processed. The obtained L3A data underwent two different calibration methods as the provider changed the calibration method starting from 24 January 2014. The geometric accuracy of the L3A product depends on the quality of the reference dataset (BlackBidge, 2015). Each tile represents an area of 25 km (5000 lines) by 25 km (5000 columns), with overlapping areas of 500 m on all sides. However, due to the persistent cloud cover over the study area, only 26 tiles were acquired during the requested period. Out of the 26 total tiles, 14 tiles were collected over the study area from different acquisition dates. Additionally, the received data spanned the two different calibration methods. Only three RapidEye tiles were needed to cover the Sebangau study area. These three images were acquired on September 20th, 2013 and were selected because they had the least percentage of clouds over the study area due to it being the dry season. The data were delivered with the same calibration method in UTM map projection for WGS 84 Datum Zone 49S.

Additionally, RapidEye data of L3A product over the Hainich study site, Germany, were used in this study. The Hainich study area was covered by four RapidEye tiles. The images acquired on September 29th, 2011 with no cloud cover were then selected. The data were delivered in UTM map projection for WGS 84 Datum Zone 32 North.

Chapter 3

Methods

3.1. The effect of forest restricted visibility on the basal area estimates from ACP in two different forest types of Sebangau and Ncaute

3.1.1. Estimation of the distance of visibility in Sebangau forest and simulation of ACP trees

This study determined the distance of visibility for ACP in a dense peat-swamp forest with restricted visibility as one of the criteria to identify a suitable BAF that minimizes observation error as a function of 'level of visibility'. The distance of visibility in this study characterized the distance to identify and count trees in the field using ACP. The complete field measurements of tree *DBH* and position in the FC plot enabled the development of the tree map shown in Fig. 2.5, and a simulation of the ACP trees in an experimental setting. Additionally, the study compared BA estimates derived from ACP and FAP in both the Sebangau forest (limited visibility) and the Ncaute forest (good visibility) to improve the understanding of visibility effects in forests with different structures.

When simulating ACP trees, all reference points lying close to the plot border were excluded so as to avoid edge effects (Fig. 2.5). Thus, $n = 37$ out of 64 reference points were used. The distance of visibility was assessed for each reference point by assuming that the total number of trees counted (z) in the field using BAF = 1 were the z nearest trees to the reference point on the tree map, thereby assuming similar visibility in all directions. The distance between the reference point and the z^{th} tree corresponded to the distance of visibility at that reference point. However, the "true" distance of visibility actually lies at an indeterminable point between the z^{th} and the $z+1^{th}$ trees. In this study, the maximum distance of visibility for the entire FC plot was defined as the mean distance from all reference points ($n = 37$), taking into account the standard deviation of visibility distance from all reference points.

In theory, a tree is counted into an ACP sample of a sample point if the sample point lies within the tree circular inclusion zone describing its probability of selection. All such trees are theoretically to be counted, hereafter referred to as theoretically-counted tree. To determine whether a sample point lay within the inclusion zone of a given tree, the radius of the circular

inclusion zone of the tree (R_t) and the horizontal distance of the tree from the sample point were compared. R_t was calculated as a factor of the employed BAF and the dimension of the tree (*DBH)* following (Bitterlich, 1984),

$$R_t = \frac{50}{\sqrt{BAF}} \cdot DBH$$

For each reference point, the theoretically-counted trees were simulated using different BAFs from 2 to 10, sequentially. The distance between the reference point and the z^{th} theoretically-counted tree corresponded to the required distance of visibility. Among these theoretically-counted trees were trees located within the distance of visibility that had been determined for each reference point, referred to as field-counted trees. Whenever the number of field-counted trees was smaller than the number of theoretically-counted trees, it indicated that some of the trees that theoretically should be counted were not counted during field measurements. These trees were referred to as field-unseen trees. This study assumed that any field-unseen trees were omitted due to the restricted visibility of the reference point. Even though the reference point lay within the circular inclusion zone of the trees, their positions were farther than the distance of visibility. However, this is a strong assumption because other explanations could underline the difference between theoretically- and field-counted trees, such as measurement error or miscounting of border trees. By assuming that the difference between theoretically- and field-counted trees was only due to visibility, this study potentially overestimated the effect of restricted visibility on basal area estimates. All simulations were done using R software (R Core Team, 2015).

Fig. 3.1 illustrates the distance of visibility d_v for a reference point. The field-counted, theoretically-counted, and field-unseen trees that were simulated for different BAFs are also shown. In the example, there should have been 11 theoretically-counted trees. Due to the d_v of the reference point, only 6 trees were counted in the field. It is assumed that the 7^{th} to 11^{th} trees were the field-unseen trees as a result of restricted visibility from the sample point. The general workflow of the study is shown in Fig. 3.2.

3.1.2. Selection of a suitable BAF for the Sebangau forest

In order to determine the suitable BAF for the forest, the following criteria were tested for each BAF: (i) the conformity between the required distance of visibility and the mean distance of visibility for the entire FC plot, (ii) the difference between basal area estimates from the theoretically- and the field-counted trees, tested with the paired sample *t*-test ($p = 0.05$), and

(iii) the average number of field-counted trees. By considering these criteria, the suitable BAF should be large enough that the distance of the z^{th} theoretically-counted tree will not exceed the maximum distance of visibility. Additionally, there should be no significant difference between basal area estimate derived from the theoretically- and field-counted trees. Also, the BAF should not be so small that it restricts the number of sample trees counted in each sample point. An average number of trees between 5 to 16 per sample point were considered suitable in this study, following the recommendation of earlier studies: between 5 to 15 trees (Avery and Burkhart, 2002), 6 to 16 trees (Van Laar and Akça, 2007), 10 to 15 trees (West, 2009), or at least 10 trees (Bitterlich, 1984).

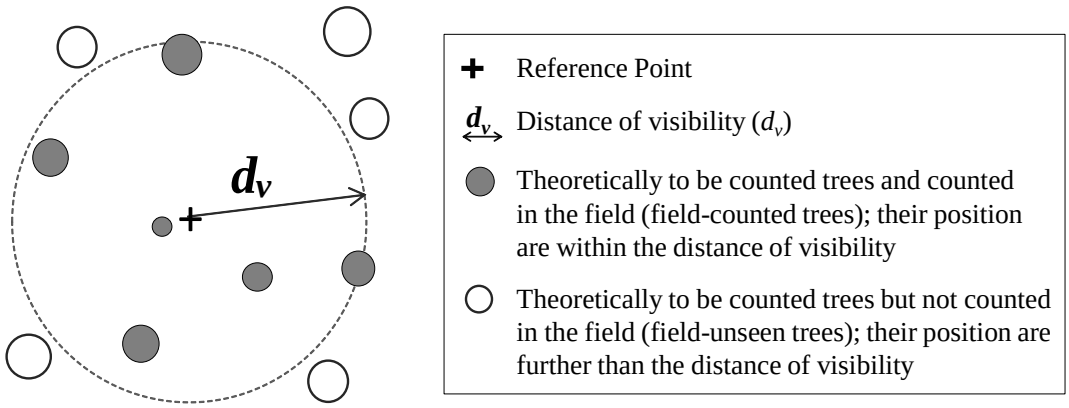


Fig. 3.1. Determination of distance of visibility (d_v) as the distance from the sample point to the farthest field-counted tree; the field-unseen trees (white circles) are left out.

3.1.3. Comparison of basal area estimates of ACP and FAP in Sebangau and Ncaute forests

A comparative study was conducted for a better understanding of the effect of forest restricted visibility in basal area estimates from ACP. The basal area estimates from ACP of the closed forest of Sebangau were compared to the basal area estimates of the open savanna forest of Ncaute. The input datasets were the basal area estimates from ACP and FAP of both study areas. In case of Sebangau, only the LSI plots in which ACP was conducted for BAF 1, 2, and 4 were used (n = 21). For Ncaute datasets, data from n = 33 plots were used since basal area from ACP were estimated with BAF 1, 2, and 4 in all plots. Assuming the absence of measurement errors, both ACP and FAP allow for unbiased estimation of basal area. As both plot designs were installed around the same sample points, these paired datasets of ACP and FAP produced dependent pairs of stand basal area estimates. There is no statistically significant difference between the two estimations, because they both estimate the same population. The reference used the basal area estimated from FAP, since the population

parameter was not available. The pairs of per-plot basal area estimates from FAP and ACP were tested with the paired *t*-test ($p = 0.05$) using the *t*.test() function of R (R Core Team, 2015). Prior to the test, normality of the data was assessed graphically using a Q-Q plot. Comparisons between the two approaches of different BAFs (1, 2, and 4) were conducted to assess the consequences of selecting a BAF on the basal area estimates from ACP between a forest with limited visibility and a forest with no visibility restriction.

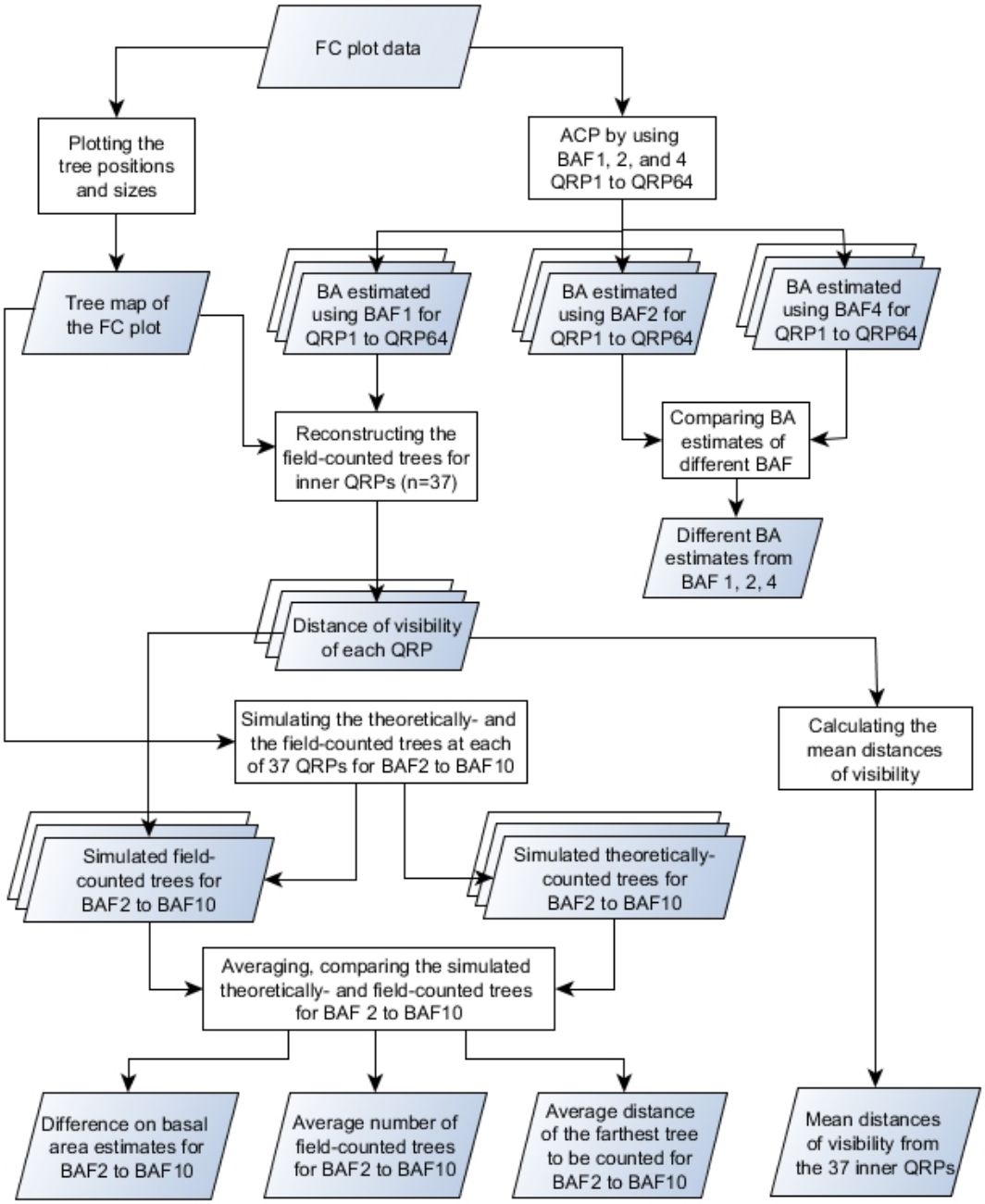


Fig. 3.2. General workflow for estimating the maximum distance of visibility by using the full census (FC) data. ACP for Angle Count Plot, BA for Basal Area, BAF for Basal Area Factor, QRP for Quadrant-Reference-Point; the point where ACP was installed.

3.2. Integration of field inventory and optical RapidEye data for the AGB estimation models in two different forest types of Sebangau and Hainich

3.2.1. Allometric model to calculate tree AGB

Selection of a suitable biomass allometric model is essential (Pérez-Cruzado et al., 2015; Picard et al., 2016). Different models may produce considerably different AGB estimations for the same set of sample trees (Skole et al., 2011). For example, a study of Rutishauser et al. (2013) accounted for an absolute difference of 10-15% on average between *DBH* models and models including height measurement. Individual tree AGB in the Sebangau dataset was estimated using the multi-species model of Manuri et al. (2014) with *DBH*, wood density (*WD*), and height (*H*) as input variables:

$$Estimated\ AGB = 0.15 * DBH^{2.095} * WD^{0.664} * H^{0.552}$$

where AGB in Mg ha^{-1}, *DBH* is in cm, *WD* in g cm^{-3}, and *H* in m.

The model from Manuri et al. (2014) was considered to be the most appropriate for AGB estimation in the Sebangau peat-swamp forest based on following characteristics. The model was developed by felling and determining the AGB of 148 sample trees (24 taxonomical families) from Indonesian peat-swamp forests in Riau and South Sumatera (Sumatera) and in West Kalimantan (Kalimantan) (Manuri et al., 2014). The 148 sample trees had a *DBH* range of 2-167 cm, a wood density range of 0.242-1.000 g cm^{-3}, and a tree height range of 2.8-49.5 m. These ranges covered the tree measurements from the Sebangau dataset. Information for the input wood density variable was taken from the list of tree species wood density of Sebangau that was compiled by the project KL894/17. In cases where the tree species could not be identified, the wood density for such tree was calculated as the average wood density of the 10 most abundant tree species in the dataset by terms of the basal area (Fischer et al., 2011). The heights of the non-height-measured trees were predicted based on the 10 tree-height measurements per plot. The models were built separately for each plot under the project of KL894/17. However, this modelling was not part of the study, but rather the output was used.

For the Hainich dataset, tree AGB was calculated by considering two general tree types: deciduous and coniferous. The multi-species allometric models of Fehrmann (2007) were used, with *DBH* and height (*H*) as the input variables:

$$Estimated\ AGB\ (Broadleaf) = 0.0349 * DBH^{2.0476} * H^{0.7589}$$
$$Estimated\ AGB\ (Conifers) = 0.0818 * DBH^{2.1704} * H^{0.3492}$$

where *DBH* is in cm, *H* in m, and AGB in Mg ha^{-1}.

The allometric models were developed using sample trees mainly from temperate forests in Germany and Central Europe. The broad-leaved tree model of Fehrmann (2007) was built from 528 sample trees of 8 different species with a range of *DBH* between 1 and 77.1 cm. The coniferous tree model was built from 963 sample trees of 7 different species with a *DBH* range between 2 and 73.8 cm. These *DBH* ranges covered more than 99% of the tree measurements in the EP and the AWF datasets, thus considered suitable for the current study. As in the Sebangau, biomass models in Hainich required *DBH* and *H*, while *H* was not measured in the field for all sample trees. Therefore, prior to the AGB calculation, a general logarithmic model (Schmidt, 1967; Van Laar and Akça, 2007) was selected to fit a *DBH* – *H* curve, using the available tree height measurements as input data. The resulting model was then used to generate the missing tree height for all non-*H*-measured trees.

3.2.2. AGB estimation based on field samples

AGB at the plot level was calculated by aggregating all individual sample tree biomass estimations per plot. For the concentric plot design, the aggregation was done after applying an expansion factor for each sub-plot to transform the plot-wise measurement into a per hectare estimation.

Simple random sampling (SRS) estimators were used to estimate the AGB of the study area, including its corresponding variance and standard error. This SRS estimator framework is frequently used to analyze systematic samples in the absence of unbiased statistical estimators for systematic sampling. Applying the SRS estimator to systematic samples is known to be unbiased at estimating the mean but conservative at estimating the standard error (Köhl et al., 2006). The field-based estimated mean AGB $\bar{y}$ (Mg ha^{-1}), the estimated variance s^2, and the estimated standard error of the mean SE were then calculated as follows:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}$$

$$SE_{(\bar{y})} = \sqrt{\frac{s^2}{n}} = \sqrt{\hat{\mathrm{var}}_{(\bar{y})}}$$

where y_i is the estimated AGB for plot *i*, *n* is the sample size, and $\hat{\mathrm{var}}_{(\bar{y})}$ is the estimated variance of the estimated mean AGB.

3.2.3. Pre-processing of RapidEye data

All spectral bands of RapidEye tiles were converted to the top of atmospheric (TOA) reflectances according to BlackBidge (2015). Furthermore, an atmospheric correction was applied using the 6S (Second Simulation of Satellite Signal in the Solar Spectrum) algorithm (Vermote et al., 2006) within the ForestEye Processor (Magdon et al., 2011). The standard method of the 6S algorithm was selected and had the following input parameters: designation of aerosol's model, atmospheric profile, and meteorological visibility. Considering the geographical location of the study areas, the maritime aerosol model with tropical atmospheric profile was deemed appropriate and selected for the Sebangau. The meteorological visibility was calculated from the Aerosol Optical Thickness (AOT, τ_{550}) at 500 nm following Bojanowski (2007). The AOT value was obtained from the closest available station of the Aerosol Robotic Network (AERONET, https://aeronet.gsfc.nasa.gov/) of the National Aeronautics and Space Administration (NASA). The corresponding station was the Palangka Raya station located about 15 km from the Sebangau study site. The AOT value closest to the acquisition time of the images was selected to represent the atmospheric conditions at the time the images were acquired. The atmospherically corrected tiles were then mosaicked and masked to the study area boundary.

A striping problem was observed in the blue band of the Sebangau data most likely resulting from the image calibration process (Pfeifer et al., 2016). Hence, this band was dropped and excluded from further processing and analysis steps. Clouds, cloud shadows, and haze areas were visually identified and delineated, before being masked out from the final data.

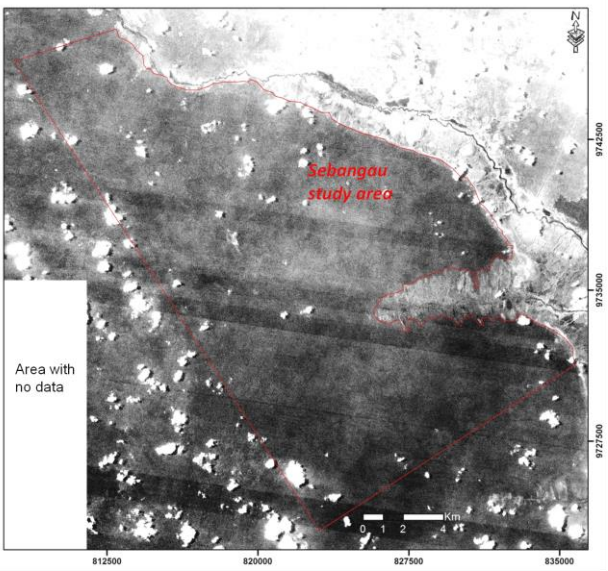


Fig. 3.3. Striping problems of RapidEye blue band (band 1) over the Sebangau study area (red line). The image acquired on 20th September 2013, presented in gray scale.

3.2.4. Deriving potential predictor variables from RapidEye data

Various predictors for AGB can be derived from remote sensing data. This study assessed three types of predictor variables (Table 3.1) which were calculated using all RapidEye bands except the blue band. These variables included the reflectance values of the original bands, vegetation indices, and texture features from the grey-level co-occurrence matrix (GLCM).

Table 3.1. Predictor variables derived from RapidEye images.

Independent Variables	**Description/Equation**	**References**
Spectral reflectance of RapidEye Band 2 to Band 5	Green (520-590 nm), Red (630 - 685 nm), Red-edge (690-730 nm), Near Infrared (760-850 nm)	(BlackBidge, 2015)
Vegetation indices		
Normalized Difference Vegetation Index (NDVI)	$\frac{B5-B3}{B5+B3}$	(Rouse et al., 1974)
Normalized Difference Vegetation Index Red-Edge (NDVIRE)	$\frac{B5-B4}{B5+B4}$	(Gitelson and Merzlyak, 1997)
Normalized Difference Vegetation Index Red-Edge (NDRE)	$\frac{B4-B3}{B4+B3}$	(Gitelson and Merzlyak, 1994)
Canopy Chlorophyll Content Index (CCCI)	$\frac{(B5-B4)/(B5+B4)}{(B5-B3)/(B5+B3)}$	(Barnes et al., 2000)
Modified Soil Adjusted Vegetation Index (MSAVI)	$\frac{2*B5+1-\sqrt{(2B5+1)^2-8(B5-B3)}}{2}$	(Qi et al., 1994)
GLCM Texture features		
Mean (MEAN)	$\sum_{i,j=0}^{N-1} i\mathrm{P}_{i,j}$	(Haralick et al., 1973)
Homogeneity (HOMO)	$\sum_{i,j=0}^{N-1} i\frac{\mathrm{P}_{i,j}}{1+(i-j)^2}$	(Haralick et al., 1973)
Entropy (ENTR)	$\sum_{i,j=0}^{N-1} i\mathrm{P}_{i,j}(-\ln \mathrm{P}_{i,j})$	(Haralick et al., 1973)
Correlation (CORR)	$\sum_{i,j=0}^{N-1} i\mathrm{P}_{i,j}\left[\frac{(i-\mu)(j-\mu)}{\sqrt{\sigma_i\sigma_j}}\right]$	(Haralick et al., 1973)

P is the dimension of the GLCM, *i* and *j* as the i^{th} and j^{th} element of P. P*ij* represents the frequency of the value of reference pixel *i* in combination with the adjacent pixel *j*. P is the frequency of the cell row and column *i* and *j*; μ is the Mean and σ is the Variance.

Each original spectral band contains specific information on a viewed object. However, two or more spectral bands are usually needed to characterize the object. For example, vegetation indices (VIs) are numerical combinations of two or more original spectral bands (Huete et al., 2002) which are developed to achieve better information related to biophysical vegetation parameters, e.g., leaf area index, biomass, compared to any single original band (Jones and Vaughan, 2010). The most common combination used for creating VIs is the red and near-infrared (NIR) spectra of the electromagnetic spectrum (Myneni et al., 1995). As depicted in Fig. 2.10, the difference between the reflectance in the red and NIR spectral regions creates a variable that is sensitive in detecting the presence of green vegetation (Bannari et al., 1995). The use of VIs also reduces the confounding effects of soil background, illumination angle, canopy geometry, and atmospheric conditions in the canopy reflectance of the original spectral bands (Jackson and Huete, 1991; Jones and Vaughan, 2010; Myneni et al., 1995).

Five VIs were computed in this study (Table 3.1). The first was the most widely used VI, the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974). The next three were selected in combination with the red-edge band since they were expected to be more sensitive to the variation of the chlorophyll content in the vegetation. These are Normalized Difference Vegetation Index Red-edge (NDRE), Normalized Difference Vegetation Index Red-edge (NDVIRE), and Canopy Chlorophyll Content (CCCI). The last index was the Modified Soil Adjusted Vegetation Index (MSAVI), an index that consider the variation from soil backgrounds in low vegetation cover and is suitable for vegetation sensitivity as well as soil noise reduction (Qi et al., 1994). Moreover, the index has been successfully selected as one of the predictor variables for AGB estimation model in Borneo's tropical forests (Pfeifer et al., 2016).

Information about an object or a land cover type in remote sensing data can generally be identified by shape, size, tone or color, pattern, shadow, and texture of the object. Image texture contains information about the spatial arrangement of tonal variation (Eichkitz et al., 2015). Some objects can be identified more easily by using textural features rather than by the spectral reflectance properties only (Haralick et al., 1973). One of the most common approaches to extract textures from remotely sensed data is the GLCM texture features of Haralick et al. (1973). These GLCM textures measure the frequency of grey-tone levels of adjacent pixels within a define window size. Previous studies showed that the inclusion of GLCM texture features as predictor variables improved AGB model performance (Eckert, 2012; Lu, 2005; Sarker and Nichol, 2011; Wijaya et al., 2010a). However, not all texture measures are useful for extracting AGB information.

Selection of a suitable texture measure and the associated window size for extracting AGB information also depends on the characteristics of the forest, the remote sensing data, and the band of the image. There is no standard approach in selecting proper texture features for estimating biomass since textures vary according to the studied landscape and the type of remote sensing data (Lu, 2005). In general, the GLCM texture features can be grouped into three main categories: orderliness, contrast, and descriptive statistics (Hall-Beyer, 2007). The texture features within one category contain similar information that tends to be highly correlated when used together, particularly if derived from the same band. Hence, the simultaneous use of all types of texture features as independent variables for a regression model may cause multicollinearity issues. Additionally, such simultaneous use may not be efficient since the computation of GLCM textures is also time-consuming (Hall-Beyer, 2007), especially in large area of interest. For these reasons, one type of GLCM texture feature was selected from each category. In addition, the GLCM texture mean was selected since it correlated the least with the other GLCM texture features. Therefore, in total four types of GLCM texture features were selected for this study. All texture features were derived from the red-edge band using the 'raster' package (Hijmans, 2016) and the 'glcm' package (Zvoleff, 2016) implemented in R software (R Core Team, 2015).

Image texture also depends on the window size from which the texture features are extracted. During the pre-processing step, texture features derived from the window size of 7x7, 11x11, and 15x15 were examined. Through visual investigations, texture features derived from the 3x3 and 5x5 window sizes contained large inter-pixel variations and were therefore considered too small. Window sizes larger than 15x15 were considered too large, as they smoothens the variability among pixels and possibly lose important variation in pixel brightness. The correlation magnitude r between the field-based AGB estimate and the texture-features derived from the window size of 7x7, 11x11, and 15x15 ranged from 0.45-0.49, 0.49-0.51, and 0.51-0.52, respectively. Thus, the GLCM texture features computed from the 15x15 window size that gave stronger correlations between field-estimated AGB and texture-predicted AGB were selected as predictor variables. To overcome the edge effects on the derived texture features, an additional buffer of 500 m was added from the boundary of the study area. In this step, a total of 40 potential predictor variables derived from RapidEye data were available.

3.2.5. Extracting information from remotely sensed predictor variables at plot level

Using a commonly adopted method, the information from remotely sensed AGB predictor variables were extracted at the plot level as the mean value of all pixels within the plot. In

addition to the mean value, three other descriptive statistics were also calculated to characterize each plot including: maximum, minimum, and standard deviation. Pixels to be included in the calculation were restricted to the ones in which the center was located at or within the plot boundary. In case of a circular plot like in the Sebangau, the pixels included in the calculation were the ones in which the center was located at an equal or lesser distance than the plot radius (m), as illustrated in Fig. 3.4.

Since the information from remotely sensed data was extracted according to the geographical position of the plots, it was important to analyze the mismatch of positions between the remote sensing data and the ground coordinates (Gobakken and Næsset, 2009). An initial investigation of the mismatch position was conducted to identify a suitable area size (plot size) for extracting the remote sensing information. In total, 18 ground control points (GCP) were recorded. However, attaining precise GCPs inside the Sebangau forest was difficult because the closed canopy of the forest likely degraded the received and the reflected signals from the GPS device. Therefore, the GCPs were instead collected in Palangka Raya city. I assumed the same location shift between the field GPS measurements and the RapidEye image coordinates over the Sebangau forest and the city area. Co-registration between the GCPs and the RapidEye image showed an *RMSE* of 0.77 pixels. Thus, a 4 m buffer distance was added to the 16 m radius of the LSI plots to counter the mismatch. Similarly, a 2.4 m buffer distance was added to the 12.6 m radius of the AWF-Hainich plots. The additional 2.4 m buffer distance was determined both by considering the higher accuracy of the RapidEye image over the Hainich forest compared to that of the Sebangau forest (BlackBidge, 2015) and also to have a convenient number.

Information from all remotely sensed predictor variables was then extracted from a 20 m radius plot size for the Sebangau dataset, a 15 m radius plot size for the AWF-Hainich dataset, and a 1 ha plot size for the EP-Hainich dataset. For each dataset, the mean, minimum, maximum, and standard deviation were calculated for each predictor variable from the original band reflectances and vegetation indices. In addition, mean values were calculated from the texture features. The values were then assigned to the plot centers.

For the Sebangau study, information from all predictor variables was extracted on the systematic plots of LSI (n = 46). Three plots were located under clouds or cloud shadows, so that no remote sensing data were available for these plots. Therefore, the final number of plots used for establishing the AGB model in the Sebangau study was n = 43. For the Hainich study, information from all predictor variables was extracted on both the EP (n = 28) and the AWF

(n = 177) datasets. Each of the datasets was independently used to establish an AGB estimation model for Hainich.

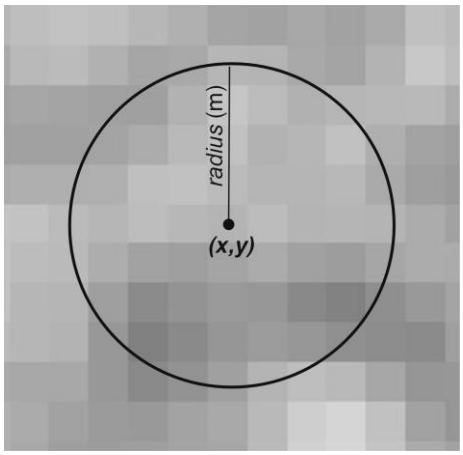


Fig. 3.4. Plot boundary (black circle) of a defined radius (m), and a plot center point (x, y), overlaid on the 5 m pixel size of a RapidEye predictor variable. Information from the predictor variable was extracted as the mean, minimum, maximum, and standard deviation of all pixels within the plot, and assigned to the plot center point. Pixels included in the calculation are restricted to the ones whose center point is located at or within the plot boundary.

3.2.6. Biomass modeling and variable selection approach

In this step, AGB was linked to remotely sensed predictor variables. A Box-Cox power transformation (Box and Cox, 1964) was applied to all predictor variables to test the normality and linearity of the variables using the 'MASS' R package (Venables and Ripley, 2002). The relationship between the field-based AGB and each predictor variable was evaluated using the Pearson correlation coefficient r and the corresponding p-value. The correlation was calculated by using the 'Hmisc' R package (Harrell Jr et al., 2016).

In the modelling approach, field-based AGB was the response variable and remotely sensed variables were the potential predictors. A multiple linear regression analysis with a stepwise (exhaustive) variable selection procedure was conducted to fit the AGB model:

$$AGB = ß_0 + ß_1 * X_1 + ß_2 * X_2 + ... + ß_i * X_i + \varepsilon_i$$

where $ß_0$ is the model intercept, $ß_1$, $ß_2$, .., $ß_i$ are the regression coefficients for the predictor variables 1, 2,..i, X_1, X_2, ...X_i are the predictor variables 1, 2,..i, and ε_i is the random error term.

The variable selection was conducted using the 'leaps' R package (Miller, 2009), where the regression algorithm was applied for any combination of two to eight predictor variables. For

each model combination, the Adjusted R^2 (Adj R^2), the residual standard error (RSE), and the statistical significance of each selected predictor variable were investigated. In order to produce a more stable model and to avoid overfitting, a supplementary analysis to measure the magnitude of multicollinearity among the model predictors was applied. The variance inflation factor (VIF) in the 'car' R package (Fox and Weisberg, 2011) was used to assess the potential correlation between predictor variables. The VIF of a predictor *i* (VIF_i) was calculated as follows:

$$\mathrm{VIF}_i = \frac{1}{1 - R_i^2}$$

where R_i^2 is a collinearity measure between the *i*th predictor and the remaining predictors. The value represents the proportion of variance in predictor *i* that is associated with other predictors in the model (O'brien, 2007). The study of Sileshi (2014) showed that a VIF > 5 indicates significant collinearity. To minimize the multicollinearity effects on the model, a cut-off VIF of ≤ 5 was employed, as recommended by O'brien (2007). The final AGB model was then defined as the one that provided the highest Adj R^2, the lowest RSE, and each selected predictor variable in the model has a statistical significance of a *p*-value ≤ 0.05 and a VIF ≤ 5.

The performance of the final AGB model was assessed with the leave one out cross validation (LOOCV) approach. The approach calls for the calculation of the cross-validated root mean squared error (*RMSE*) between observed and predicted AGB. *RMSE* is a good measure to assess how accurately a model predicts its target variable; hence it is considered the most important criterion for the purpose of model prediction (Yoo et al., 2012). In the LOOCV approach, the model was fit for *n* iterations. In each iteration, the model was trained from the *n*-1 plots and tested to predict the AGB of the excluded plot. Each time, a random plot was excluded and used as the tested case. For each iteration, the residuals between observed and predicted AGB were calculated for the excluded plot, including the *RMSE*. The final *RMSE* was calculated as the average *RMSE* value over all iterations. The relative *RMSE* ($RMSE_r$) was calculated as the ratio between the *RMSE* and the estimated population mean from the field observation data. This procedure was done for all elements of the dataset using the 'boot' R package (Canty and Ripley, 2016).

3.2.7. Regionalization of the AGB model and model-assisted estimation

Once the final AGB model was constructed, the model was applied to the entire study area. A typical approach for spatial extrapolation of the empirical model, from training area to the

entire area of interest (AOI), is by aggregating pixels of the remote sensing data into grid cells of similar spatial scale as the area used for extracting the remotely sensed predictor variables. An aggregation level of 35 m x 35 m was used for the Sebangau study, which corresponded to the field plot size of the area used for extracting the RapidEye information. In the Hainich study, an aggregation level of 100 m x 100 m was used for the EPs, and an aggregation level of 25 m x 25 m was used for the AWF plots. Because the population was considered to be the entire AOI, each grid cell represented one population element, for which AGB was predicted.

Since this study used probability-based sample plots, the model-assisted regression estimator was used for calculating the mean AGB per hectare of the entire study area $\hat{\mu}_{MA}$ (Mg ha^{-1}) according to McRoberts et al. (2013a),

$$\hat{\mu}_{MA} = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i + \frac{1}{n}\sum_{i=1}^{n}\hat{e}_i$$

where $\hat{y}_i$ is the AGB per hectare of the population element *i* predicted by the model, N is the population size, and $\hat{e}_i$ is the model prediction residual for plot *i* which is calculated over the sample units as the deviation between observation value and model prediction for plot *i*, $\hat{e}_i = y_i - \hat{y}_i$. In the case that the least squares parameter estimation approach for a linear model is applied, as it was in the current study, the sum of the residual will be zero. This sum can theoretically also be different from zero, for example when the model was developed using observations from outside the sample (Næsset et al., 2016).

The final AGB predicted maps of Sebangau, EP-AGB, and AWF-AGB are given with 35 m x 35 m, 100 m x 100 m and 25 m x 25 m spatial resolutions, respectively. To avoid negative biomass values, the minimum value of the predicted AGB was set to 0 with a consequence of potentially overestimated AGB of the study areas.

3.2.8. Independent validation of the AGB predicted map

An independent validation of the Sebangau AGB predicted map was conducted against the AGB estimates from HIL plots (*n* = 27). Eight HIL plots had to be excluded due to clouds and cloud shadows. An independent validation was also conducted for each of the produced AGB predicted maps in the Hainich study. The EP-AGB predicted map was validated against the AGB estimates from AWF dataset, and vice versa. Performance of the model prediction was evaluated using the Pearson correlation coefficient and the absolute *RMSE* between the model predicted and validation data.

3.2.9. Efficiency of RapidEye-assisted AGB estimation

The aim of the study was to quantify the improved precision of AGB estimation from the inclusion of RapidEye data compared to the precision achieved strictly from observation estimates. The magnitude of this improvement was assessed according to the relative efficiency (*RE*) between the two approaches (Mauya et al., 2015; McRoberts et al., 2016; Næsset et al., 2016). The *RE* measured the estimated variance of AGB estimates from the RapidEye-assisted method, relative to the estimated variance from the field observation estimates only. The variance of the field-based estimate $\hat{V}_{(\bar{y})}$ was estimated according to the SRS estimator.

$$\hat{V}_{(\bar{y})} = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n\,(n-1)}$$

where y_i is the estimated AGB for plot *i*, $\bar{y}$ is the field-based estimated mean AGB, and *n* is the sample size.

The variance for the RapidEye-assisted estimate $\hat{V}_{(\hat{\mu}_{MA})}$ was estimated according to the McRoberts et al. (2013a).

$$\hat{V}_{(\hat{\mu}_{MA})} = \frac{\sum_{i=1}^{n} (\hat{e}_i - \bar{e})^2}{n\,(n-1)}$$

where $\hat{e}_i$ is the model prediction residual for plot *i*, which was calculated as $\hat{e}_i = y_i - \hat{y}_i$ and $\hat{y}_i$ is the model prediction for plot *i*, $\bar{e}$ is the mean of the residuals from all plots that was calculated as $\bar{e} = \frac{1}{n}\sum_{i=1}^{n} \hat{e}_i$. The standard error of the estimator $\hat{\mu}_{MA}$ $\left(SE_{\hat{\mu}_{MA}}\right)$ was estimated as the square root of the respective variance estimator.

And the relative efficiency (*RE*) is computed as:

$$RE = \frac{\hat{V}_{(\bar{y})}}{\hat{V}_{(\hat{\mu}_{MA})}}$$

An *RE* value of 1.3, for example, would indicate a 30% greater efficiency from the RapidEye model-assisted estimates than from the field-based estimates. Consequently, to achieve the same precision as that from the remote sensing approach requires a 30% larger sample size for the field-based estimates along with the associated additional cost (Mauya et al., 2015; Næsset et al., 2016).

3.2.10. AGB estimation model of Hainich derived from two different field inventory datasets

Analysis into the use of optical RapidEye data for estimating AGB was conducted for different inventory datasets of Hainich. An AGB-difference map was produced from the EP- and AWF-AGB predicted maps to investigate different AGB predictions between the two models.

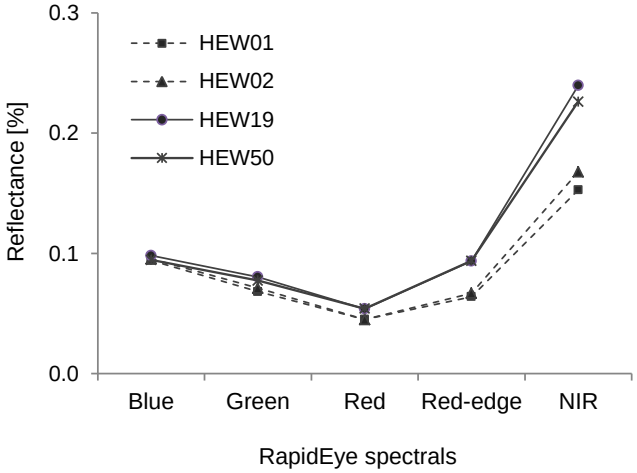


Fig. 3.5. Lower reflectances in the RapidEye spectrums from two EPs dominated by coniferous trees (HEW01 and HEW02), compared to two EPs with 100% broad-leaved trees (HEW19 and HEW50), with largest difference occurred in the NIR spectrum.

Additional AGB models for broad-leaved forest were also generated using the EP and AWF datasets, where all plots dominated by coniferous trees were excluded. In the EP dataset, two plots were dominated by over 80% of coniferous trees, whereas in the remaining plots contained less than 1%. For the AWF dataset, about 16% of the AWF plots contained higher than 1% of coniferous trees. Based on the leaf characteristics, trees in the EP and AWF datasets can be classified into broad-leaved and coniferous groups. Leaf characteristic is one of the fundamental factors influencing the spectral signatures of vegetation in optical spectral regions (Jones and Vaughan, 2010). The visible spectrum is related to the chlorophyll absorption level, and the NIR spectrum is more associated with the leaf structures (Ollinger, 2011). The plots dominated by broad-leaved trees would form a sharp contrast in the spectral signatures from the ones dominated by coniferous trees, with higher reflectance from the NIR spectrum (Fig. 3.5). The broad-leaved model was built using the reduced dataset of $n = 26$ EP plots, and $n = 148$ AWF plots. The objective was mainly to compare the resulting model performances with the one derived for the mixed forest, by using model Adj R^2, *RMSE*, and variances.

3.3. The effect of different field plot sizes on the AGB estimation models derived from field inventory and RapidEye data using the EP data of Hainich

The study was conducted to investigate the effects of different sample plot sizes in the AGB estimation in Hainich forest using the EP dataset. When field sample plots and remote sensing data are combined to make an AGB model, the use of larger plot sizes is advantageous for some specific reasons (Mauya et al., 2015; Næsset et al., 2016): (i) Larger plots increase the chance of co-registration areas between the field and the remotely sensed data. This is because the effect of plot positioning error is expected to reduce and the quality of the relationship between the two variables is expected to increase, (ii) larger plots are expected to capture more of the spatial variations between the two variables, thereby reducing the model error, (iii) larger plots are expected to minimize the effects of boundary elements since the ratio between plot perimeter and plot area decreases.

For this study, two types of datasets were derived from the EP data with a sample size of n = 26 after the exclusion of the two plots dominated by coniferous trees:

- Dataset 1

The EP square plots were gradually reduced in size by 5 m steps, from 100 m x 100 m (1 ha) to 20 m x 20 m (0.04 ha), while maintaining the plot center (Fig. 3.6a). The study focused on the effect of different plot sizes at the same sample point. In total, 17 plot sizes with n = 26 each were investigated.

- Dataset 2

Each of the EPs was subdivided into smaller sub-plots of different squares: 2x2 (50 m x 50 m), 3x3 (33.3 m x 33.3 m), 4x4 (25 m x 25 m), 5x5 (20 m x 20 m), and rectangular shapes: 1x2 (100 m x 50 m), 1x3 (100 m x 33.3 m), 1x4 (100 m x 25 m), 1x5 (100 m x 20 m), 2x3 (50 m x 33.3 m), 2x4 (50 m x 25 m), 2x5 (50 m x 20 m), 3x4 (33.3 m x 25 m), 3x5 (33.3 m x 20 m), and 4x5 (25 m x 20 m). For each dataset, one sub-plot was randomly selected from each of the EP sub-plots as shown in Fig. 3.6 (b and c).

Including the 1 ha plot data, 15 plot sizes of sample size n = 26 were considered.

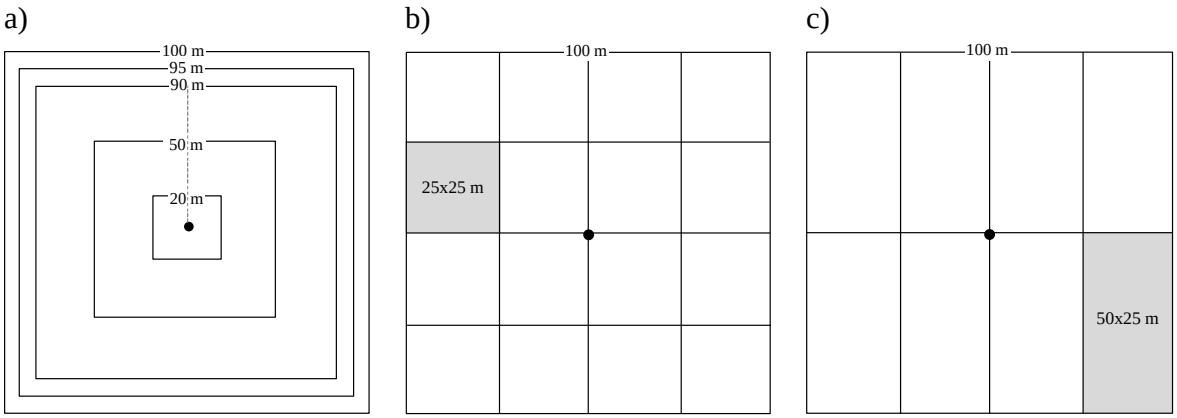


Fig. 3.6. Dataset 1 was derived by gradually reducing the 1 ha EP plot by 5 m, from 100 m x 100 m to 20 m x 20 m, while maintaining the same plot center (a). In Dataset 2, different plot sizes were derived by subdividing the EP plots into different sub-plots of regular size of squares or rectangles, in which a sub-plot was randomly selected from each of the EP sub-plots of a certain size (b, c). The gray square in b) and c) is the randomly selected sub-plot for the 25 m x 25 m and 50 m x 25 m dataset, respectively.

The effect of different plot sizes and shapes on the developed AGB model was analyzed using two approaches. The first approach involved fitting a regression for each dataset of the different plot sizes and shapes using the previously determined Hainich AGB model. In the second approach, a regression was fitted for each dataset by allowing different predictor variables to be selected in each model.

The first approach was carried out to examine the performance of the Hainich AGB model with varying plot size and shapes. In the second approach, the performance of different AGB models of different plot sizes and shapes was examined. The second approach was conducted to obtain the ‘best’ model, without imposing specific predictor variables into the model when fitting the regression of each of the datasets. Field plot level biomass estimation, modelling, and variable selection were performed as described in section 3.2. For different plot sizes and shapes, the estimated variance, standard error, and relative efficiency were calculated as described in 3.2.9. The results were compared as a function of plot sizes and shapes.

Chapter 4

Results

4.1. Major forest variables derived from field inventory plots

4.1.1. Sebangau

Table 4.1 presents the summary statistics of the Sebangau forest variables estimated from the different inventory plots, FC (n = 1), HIL (n = 35), and LSI (n = 59). According to the three designs, the estimated basal area ranged from 30.2 to 31.8 m^2ha^{-1}. The larger the sampling area, the wider the range of the basal area estimation (Fig. 4.1). The *DBH* distributions from the FC, HIL, and LSI datasets showed the typical inverse J-shape (Fig. 4.1). Twenty-five years of forest concession history has caused the present scarcity of large trees > 45 cm, with many trees < 15 cm. The higher tree density in the LSI, compared to FC and HIL, was mainly attributed to the small trees, as seen by the lower mean tree *DBH* and mean quadratic mean diameter (*QMD*) of LSI (Table 4.1). The cumulative curves in Fig. 4.1 show that 50% of the forest basal area was in the comparatively smaller trees with *DBH* sizes reaching up to 17, 18, and 15 cm, for FC, HIL, and LSI, respectively.

Table 4.1. Description of Sebangau forest derived from different sampling designs.

Variables	**FC** n = 1	**HIL** n = 35	**LSI** n = 59
Area (ha)	1	869	31,113
BA (m^2ha^{-1}) ± *SE (%)	31.5	30.2 ± 2.5	31.8 ± 3.4
Tree density (stem ha^{-1}) ± *SE (%)	2379.2	2172.7 ± 5.5	2777.8 ± 3.9
Mean *DBH* (cm) ± *SE (%)	11.1	11.8 ± 2.4	10.7 ± 1.9
Mean *QMD* (cm) ± *SE (%)	13.0	13.6 ± 2.5	12.2 ± 1.8
Number of observed tree species (unit)	114	75	87

*Applied only for HIL and LSI sampling design, QMD = Quadratic Mean Diameter.

Like in many other tropical forests, the Sebangau forest has a high diversity of tree species (Table 4.1). A total of 129 species were recorded during the fieldwork across all inventory designs, whereas 114 species were found in the FC plot. *Syzygium sp*. was the most frequent species in the FC plot, whereas *Adenanthera pavonina Linn*. was the most prevalent species in the HIL and LSI plots.

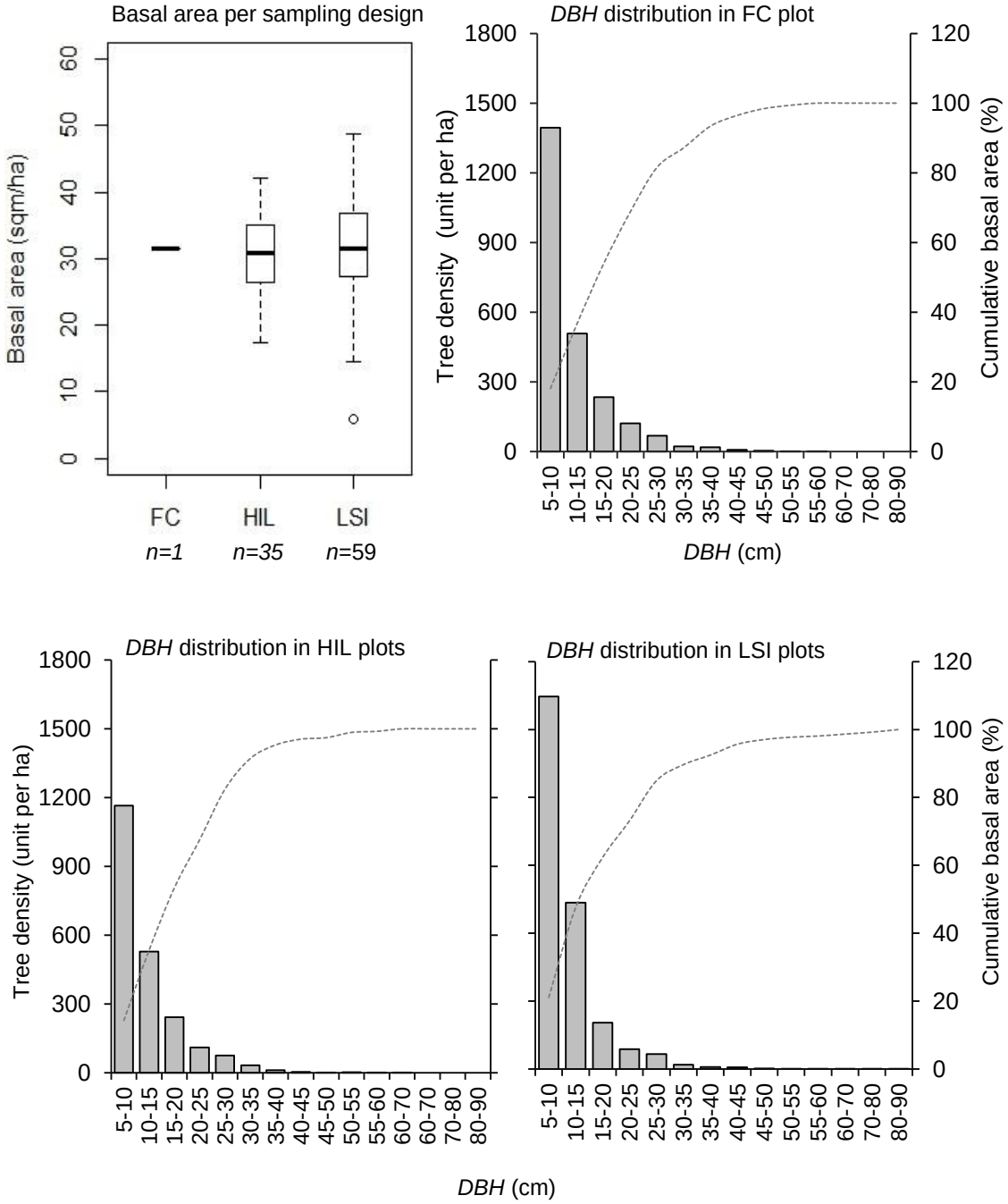


Fig. 4.1. Range of basal area estimates for FC, HIL, and LSI inventories. The *DBH* distributions were given for the FC, HIL, and LSI datasets, showing that 50% of the forest basal area was in trees of *DBH* ≤ 17, 18, and 15 cm, respectively.

4.1.2. Ncaute

Based on the field observation plots of the Ncaute study (n = 33), the estimated forest basal area was 3.6 m^2ha^{-1} with a mean *QMD* of 20.5 cm (Table 4.2). The *DBH* distributions from Ncaute followed the typical inverse J-shape with approximately 50% of the cumulative forest basal area from the trees of *DBH* < 25 cm (Fig. 4.2).

Table 4.2. Description of Ncaute forest from field sample plots.

Variables	**Ncaute** n = 33
BA (m^2ha^{-1}) ± SE (%)	3.6 ± 10.4
Tree density (stem ha^{-1}) ± SE (%)	135.5 ± 17.2
Mean *DBH* (cm) ± SE (%)	18.6 ± 5.9
Mean *QMD* (cm) ± SE (%)	20.5 ± 5.6

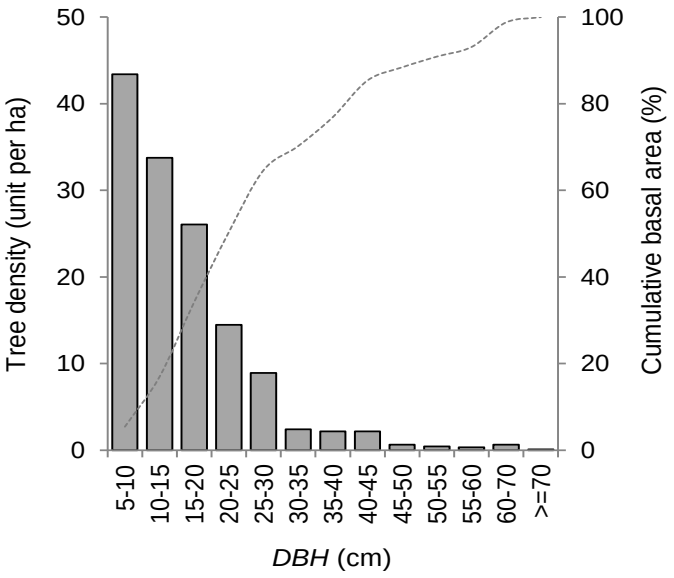


Fig. 4.2. Distribution of tree *DBH* in Ncaute plots shows that 50% of the cumulative forest basal area was in trees of *DBH* < 25 cm.

4.1.3. Hainich forest

The estimated basal area for the Hainich forest based on the EP dataset (n = 28) was greater than that based on the AWF dataset (n = 177) (Table 4.3). However, the estimated tree density from the AWF dataset was greater than that from the EP's due to the small trees from the AWF dataset, as is seen by the lower estimation of basal area and mean *QMD* from AWF. The

number of small trees in the *DBH* class of 7-10 cm and of 10-15 cm in the AWF dataset was found to be twice as high as those in the EP's (Fig. 4.3). Approximately 50% of the forest basal area was in trees of *DBH* ≤ 50 cm for the EP dataset and in trees of *DBH* ≤ 43 cm for the AWF dataset.

Table 4.3. Description of the Hainich forest from EP and AWF sample plots.

Variables	EP *n* = 28	AWF *n* = 177
BA (m^2ha^{-1}) ± SE (%)	33.9	29.4 ± 3.1
Tree density (stem ha^{-1}) ± SE (%)	388.6	527.1 ± 7.0
Mean *DBH* (cm) ± SE (%)	29.1	29.3 ± 3.4
Mean *QMD* (cm) ± SE (%)	34.4	32.4 ± 3.9

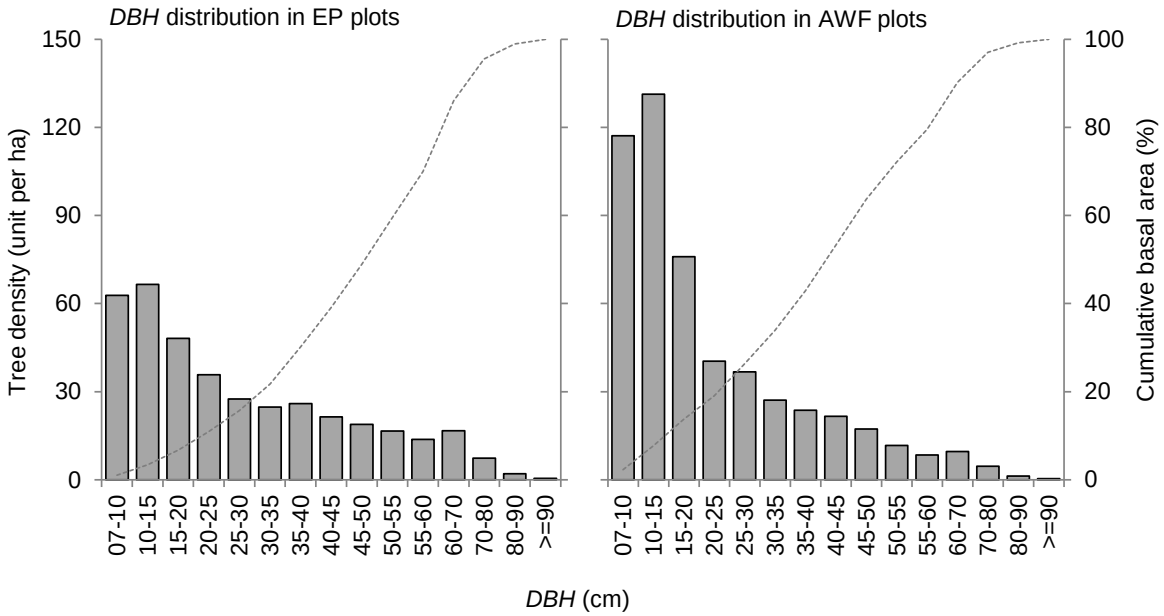


Fig. 4.3. The *DBH* distributions for the EP and AWF datasets showed that 50% of the forest basal area was in trees of *DBH* ≤ 50 and 43 cm, respectively.

4.2. The effect of forest restricted visibility on the basal area estimates from ACP in two different forest types of Sebangau and Ncaute

4.2.1. The distance of visibility for ACP in the Sebangau forest and selection of a suitable BAF

The distances of visibility over all quadrants in the Sebangau FC plot ranged between 4.5 to 13.6 m, with a mean of 8.7 m and a standard deviation of ± 2.1 m (Fig. 4.4). The defined

maximum distance of visibility for the FC plot was 6.6 m, after subtracting the standard deviation from the mean distance of visibility. The required distance of visibility for detecting the farthest theoretically-counted tree, given as the mean of all quadrants, ranked between 18 ± 3.5 m for BAF 1 to 3.5 ± 1.4 m for BAF 10.

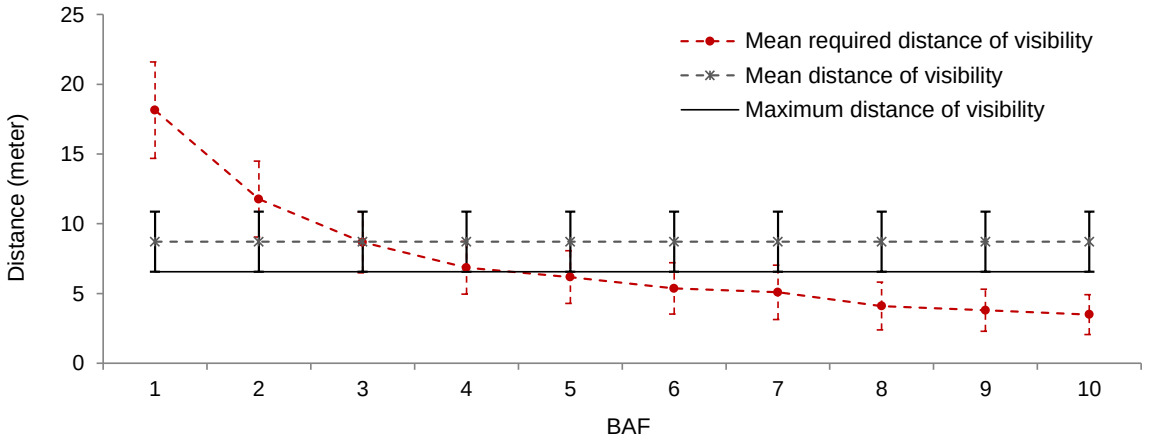


Fig. 4.4. The red-dotted line presents the mean of the required distance of visibility in the FC plot that allows the farthest theoretically-counted tree to be counted from the sample point, as a function of basal area factor (BAF). The grey-dotted line represents the mean distance of visibility of the entire plot. The empirically found "maximum distance of visibility" is presented as the black horizontal line.

Fig. 4.4 depicts conformity between the mean of the required distances of visibility with the maximum distance of visibility for a BAF of 5. When a BAF ≥ 5 is selected, the maximum distance of visibility is greater than the mean of the required distance of visibility. For BAFs < 5, the probability of omitting sample trees that should have been counted was assumed to be high. There is also conformity between the mean required distances of visibility with the mean distance of visibility for a BAF of 3. However, when a BAF of 3 or 4 is selected, the mean of the required distance of visibility is still greater than the maximum distance of visibility.

When comparing the basal area estimates derived from the number of theoretically-counted trees to the estimate derived from the number of field-counted trees with restricted visibility, selecting a BAF ≤ 6 produced significant differences in the estimation (paired *t*-test, $df = 36$, $p < 0.05$), as it is shown in Fig. 4.5. There were three sample points in which the number of theoretically-counted trees was higher than the number of field-counted trees. These sample points had trees with larger inclusion zones located outside of the maximum distance of visibility.

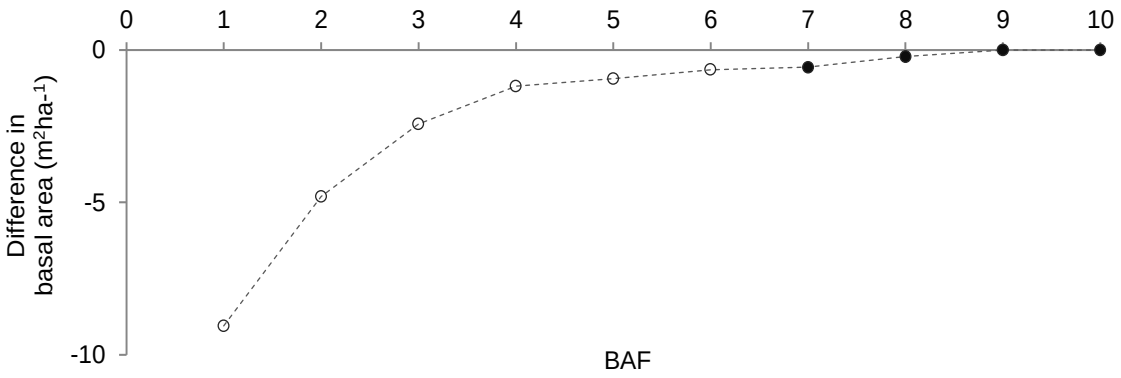


Fig. 4.5. The difference of basal area estimates between the theoretically- and the field-counted trees for different BAFs; the estimates from BAFs > 6 did not exhibit significant differences between the two estimates.

The simulation results showed that the average number of counted trees decreased from 22.5 counted trees for a BAF of 1 to 2.7 counted trees for a BAF of 10, thereby producing a basal area estimate of 22.5 m^2ha^{-1} and 26.8 m^2ha^{-1}, respectively (Fig. 4.6). With wider opening angles, smaller trees were left out, and only trees with a relatively large *DBH* were counted in the sample. By examining the number of counted trees per sample, Fig. 4.6 illustrates that BAFs from 3 to 5 included between 5 to 16 counted trees per sample point on average, as recommended by earlier studies.

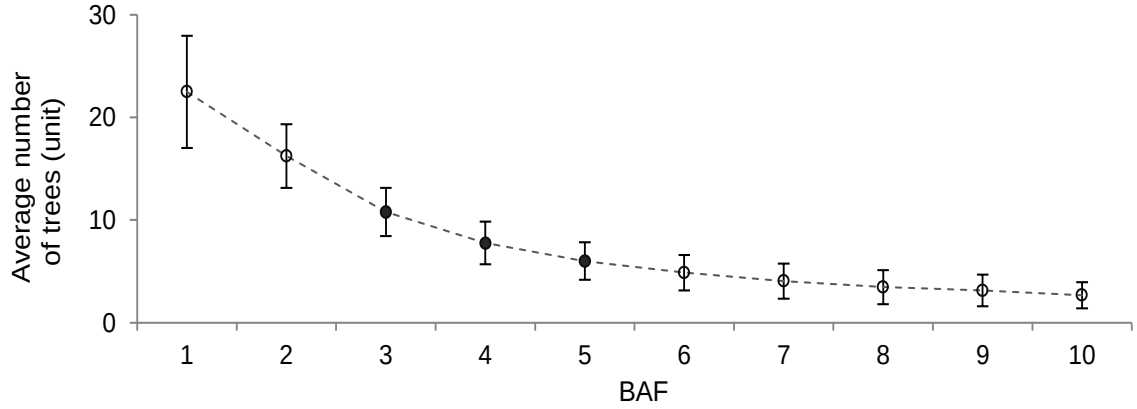


Fig. 4.6. Mean number of trees counted per ACP based on the simulation of the number of trees counted for different BAFs. BAF < 3 and BAF > 5 produced higher and lower average numbers of trees than is recommended.

Based on the aforementioned results, selection of a suitable BAF in the Sebangau forest is a trade-off between the maximum distance of visibility, the significance difference in the

estimation, and the average number of trees per sample point. However, based on the data and following the procedures applied in the current study, a BAF of 5 should be selected for implementing ACP in the dense tropical peat-swamp forest like the Sebangau with a consequence of slightly underestimating basal area. Selecting a BAF < 5 will produce greater underestimation values of basal area. When a BAF > 5 is selected, the average number of trees per sample point will be lower than the recommendation. Moreover, it will inflate the contribution of a single sample tree to the overall basal area estimate.

4.2.2. Comparison of basal area estimates from ACP and FAP datasets

In order to confirm the results in 5.2.1, a comparison between basal area estimates from ACP and FAP in the dense forest of Sebangau and in the open savanna forest of Ncaute was conducted. The same population was sampled with two approaches that allow for unbiased estimation. The mean basal area estimates from ACPs in the Sebangau forest increased monotonically over increasing BAF and approached the mean estimates of FAP (Fig. 4.7). The mean estimates from the Ncaute were relatively stable, independent of BAF (1, 2 or 4), attributed to the unrestricted visibility. These results indicate that the distance of visibility was a critical issue in the implementation of ACP in dense forests like the Sebangau, particularly when a low BAF was selected.

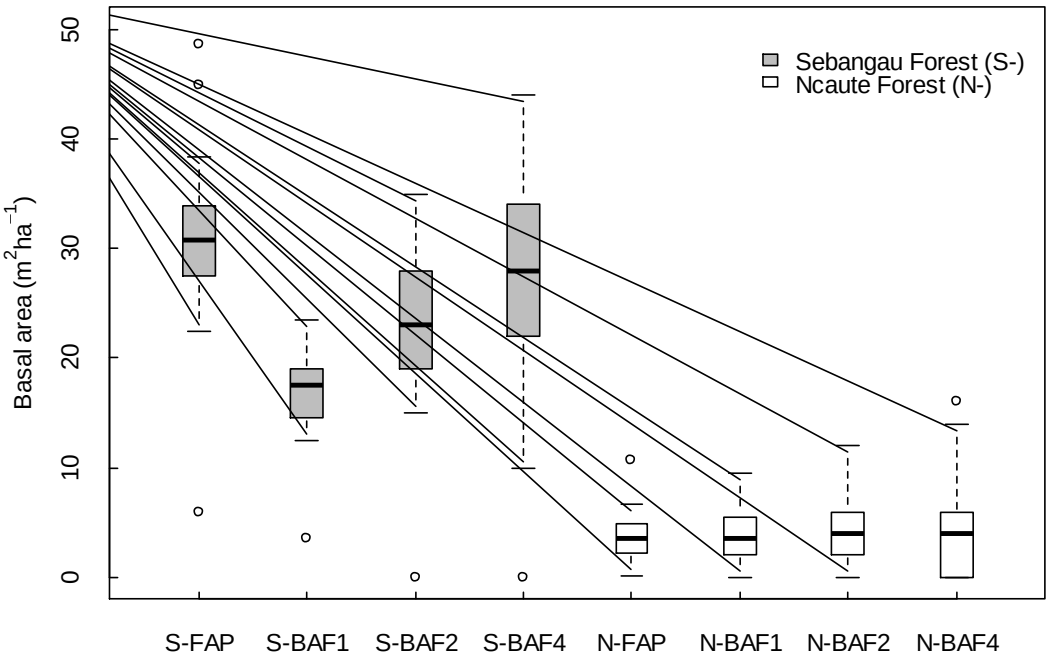


Fig. 4.7. Basal area estimates (G, m^2ha^{-1}) from FAP and ACP with different BAFs (1, 2, and 4), in Sebangau forest (S- in grey) and Ncaute forest (N- in white).

The mean basal area estimates obtained from FAP of Sebangau and of Ncaute indicated that both forests had very different basal areas (Table 4.4). The relative standard errors of the mean basal area estimates in the Sebangau study were lower compared to those in the Ncaute study. In both forests, the number of sample trees decreased with larger BAFs, while the variability of basal area estimates increased. The precision of basal area estimates decreased with larger BAFs.

Table 4.4. Estimated basal area (m^2ha^{-1}) of Sebangau (n = 21) and Ncaute (n = 33) forests from FAP and ACP; SE = Standard Error; CI = Confidence Interval at 0.95 probability level

	Mean	SE	SE (%)	CI (0.95)
Basal area (m^2ha^{-1})				
Sebangau				(df = 20)
Fixed-Area Plot	31.0	1.8	5.9	(27.2, 34.8)
Angle Count Plot				
- BAF 1	16.9	1.0	5.7	(14.9, 18.9)
- BAF 2	22.9	1.6	7.1	(19.5, 26.3)
- BAF 4	26.6	2.3	8.8	(21.7, 31.4)
Ncaute				(df = 32)
Fixed-Area Plot	3.6	0.4	10.4	(2.9, 4.4)
Angle Count Plot				
- BAF 1	3.9	0.4	11.2	(3.0, 4.8)
- BAF 2	4.2	0.6	14.8	(2.9, 5.4)
- BAF 4	4.5	0.7	15.8	(3.1, 6.0)

The mean estimate of the Sebangau study from FAP was within the confidence interval of the mean estimate from ACP with a BAF 4 (95% CI), but it was outside the confidence interval of the mean estimates from a BAF of 1 and of 2. For the Ncaute study, the mean estimate from FAP was within the 95% confidence interval of the ACP estimates from all BAFs. A paired-samples t-test was conducted to compare the mean estimates from the two approaches. There were significant deviations (paired t-test, df = 20, p = 0.05) between the mean basal area estimates from ACP and from FAP in Sebangau (Table 4.5). In Ncaute, no significant deviation between the mean estimates from ACP and from FAP was found (paired t-test, df = 32, p = 0.05). While sampling the same population indicates significant differences in observations, the differences can only be explained by observation errors. In the study of Sebangau, the difference between the two methods decreased as the BAF increased.

Table 4.5. *t*-test results comparing basal area estimates from FAP and ACP approaches in the Sebangau and Ncaute forests

		95% CI of the differerence			
	Mean	Lower	Upper	t	Sig. (2-tailed)
Sebangau (*df* =20)					
Pair 1: FAP - BAF 1	14.09	11.34	16.83	10.71	9.87E-10
Pair 2: FAP - BAF 2	8.09	5.08	11.09	5.61	1.71E-05
Pair 3: FAP - BAF 4	4.42	0.03	8.81	2.10	0.0487
Ncaute (*df* =32)					
Pair 1: FAP - BAF 1	-0.31	-1.05	0.44	-0.84	0.4061
Pair 2: FAP - BAF 2	-0.56	-1.64	0.51	-1.07	0.2923
Pair 3: FAP - BAF 4	-0.93	-2.40	0.54	-1.29	0.2070

Among the criteria to determine an optimum plot size is the number of sample trees. In this study, the mean number of trees in the FAP was higher than in the ACP (BAF 1, 2, and 4) for both forest types (Fig. 4.8). In general, the number of trees counted per ACP in the Sebangau study was 12 to 23, 8 to 18, and 3 to 11 for BAFs of 1, 2, and 4, respectively. There was one plot in the Sebangau dataset with very few trees and small *DBH*s. No tree was counted for this plot when a BAF > 1 was selected, while five trees were measured using the FAP.

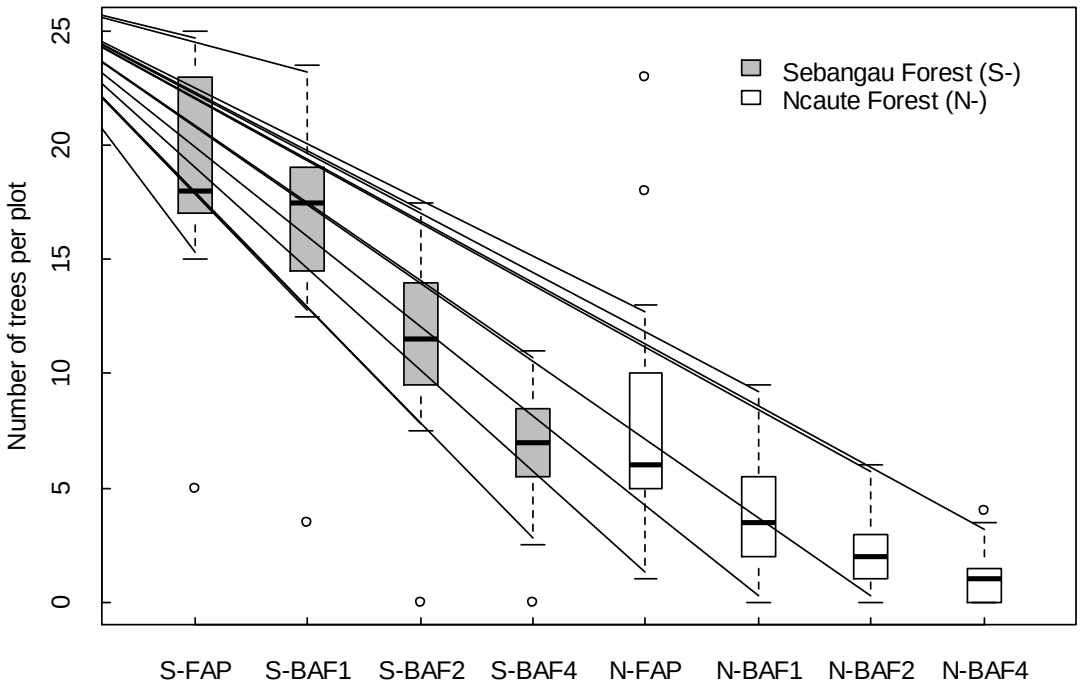


Fig. 4.8. Number of sample trees per plot from ACP of BAF 1, 2 and 4 in Sebangau forest (S- in grey) and Ncaute forest (N- in white).

In the case of Ncaute, the number of counted trees per ACP ranged from 0 to 9, 0 to 6, and 0 to 4 for BAFs of 1, 2, and 4, respectively (Fig. 4.8). This number was less than four trees per sample point, on average. For the FAP, less than four trees were measured in 18% of the Ncaute plots, while less than 8 trees were measured in 52% of the plots. Compared to the average numbers of trees per sample plot as recommended by earlier studies, the figures indicated that a BAF of 1 was already too large for an open savanna forest. The use of a smaller BAF than 1 could be more suitable for this type of forest. The argument was supported by the high standard errors of the mean basal area estimates from ACP in Ncaute, as can be seen in Table 4.4.

Results from the comparison between basal area estimates from ACP and FAP in the Sebangau and Ncaute forests confirmed the finding in 5.2.1. There was a visibility issue for estimating basal area using the ACP approach in the Sebangau forest. This study found that selecting the BAF of 5 can minimize the effect of restricted visibility on the estimation. On the other hand, a small BAF < 1 m^2ha^{-1}, should be selected for an open savanna forest like the Ncaute.

4.3. Integration of field inventory and RapidEye for AGB estimation models in two different forest types of Sebangau and Hainich

4.3.1. Relationships between field-based AGB estimates and RapidEye predictor variables

Table 4.6 presents the AGB estimates from all inventory datasets of the study areas. Based on the LSI systematic plots (n = 43), the estimated mean AGB of the Sebangau study area was 297.7 Mg ha^{-1} with a standard error of 4.4% relative to the mean (Table 4.6). A large amount of this AGB was in small trees of *DBH* < 25 cm (Fig. 4.9).

Table 4.6. Estimated AGB from field sample plots

Study area	Min	Max	Mean	SE	SE	Variance	CI (0.95)
		(Mg ha^{-1})			(%)	(Mg ha^{-1})^2	
Sebangau (n =43)	34.4	435.2	297.7	13.2	4.4	173.9	(271.1, 324.3)
Hainich - EP							
Mixed plots (n =28)	198.3	369.2	298.0	8.9	3.0	79.2	(279.7, 316.2)
Broad-leaved plots (n =26)	198.3	369.2	298.6	9.1	3.1	83.1	(279.8, 317.3)
Hainich - AWF							
Mixed plots (n =177)	1.0	575.1	244.5	8.7	3.6	76.3	(227.3, 261.8)
Broad-leaved plots (n =148)	1.0	575.2	242.3	9.4	3.9	88.6	(223.7, 260.9)

In contrast with Sebangau, large trees of *DBH* > 40 cm contributed most to the AGB of the Hainich study area, given by the EP and AWF datasets. However, different mean estimates of AGB > 50 Mg ha^{-1} were observed between the EP and AWF datasets. In comparison to the EP dataset, more trees of *DBH* < 50 cm were recorded in the AWF plots, but less trees of *DBH* > 50 cm.

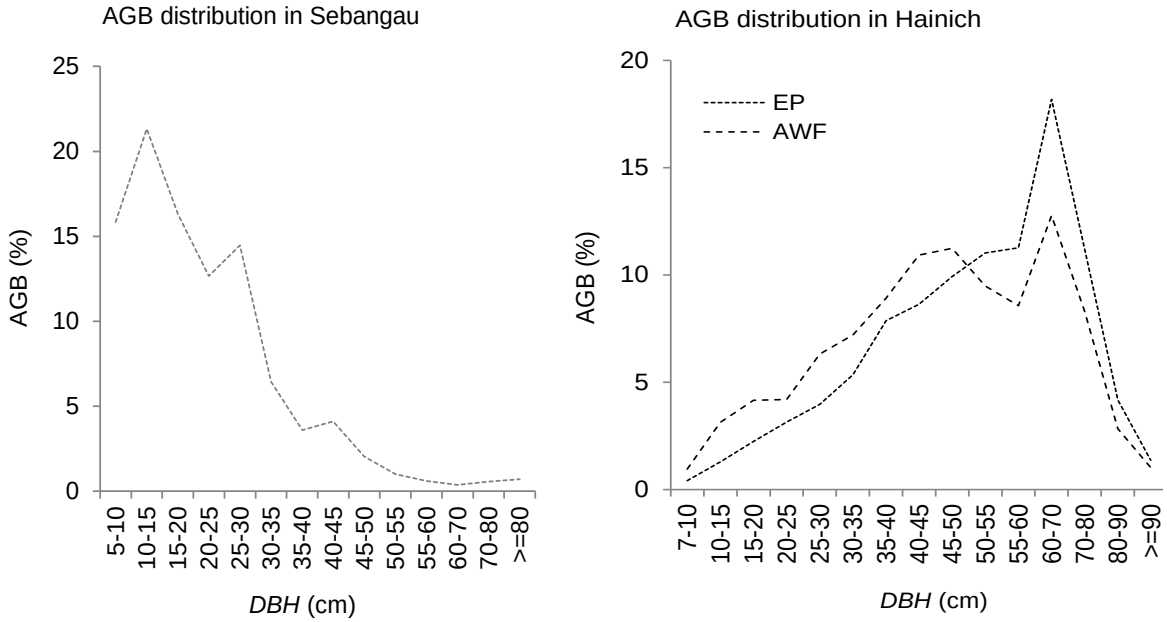


Fig. 4.9. AGB distribution over *DBH* class estimated from field sample plots in the Sebangau (n = 43), and Hainich study areas (n = 28 for EP, n = 177 for AWF).

In general, the Pearson correlation coefficients r between the plot-based AGB estimates and the RapidEye predictor variables of each dataset were low (Table 4.7). Some of these correlations were significant at $p < 0.05$ (Table 4.7, printed in bold). The predictor with the highest correlation changed for each dataset.

In the Sebangau study, only 9 out of 40 predictors were significantly correlated with the AGB (n = 43) with absolute values of r between 0.30 and 0.52 at $p < 0.05$. All these nine predictors were related to the red-edge band, indicating that red-edge is an important spectral region for AGB estimation in the dense forest of Sebangau with mixed tree species.

In reference to the Hainich study, ten predictor variables were significantly correlated with the AGB from the EP dataset (n = 28), with absolute values of r between 0.38 and 0.40 ($p < 0.05$). The exclusion of the two coniferous-dominated plots increased the number of predictors with significant correlation at $p < 0.05$ and increased the magnitude of r to 0.60.

Table 4.7. Pearson's correlation coefficients of field AGB and RapidEye predictor variables. The correlation coefficients significant at $p < 0.05$ are printed in bold; max: maximum, min: minimum, sd: standard deviation, CORR: correlation, ENTR: entropy, HOMO: homogeneity, X15 represents the window size used to calculate all texture features (15x15).

Predictors	Sebangau (*n* =43)		EP (*n* =28)		EP (*n* =26)		AWF (*n* =177)		AWF (*n* =148)	
variables	*r*	*p-value*	*r*	*p-value*	*r*	*p-value*	*r*	*p-value*	*r*	*p-value*
Band Reflectance										
B2_mean	-0.15	0.3260	0.17	0.3799	0.30	0.1414	**-0.36**	**1.23E-06**	**-0.31**	**1.46E-04**
B2_max	-0.18	0.2509	0.25	0.1985	0.26	0.2053	**-0.39**	**5.70E-08**	**-0.36**	**6.74E-06**
B2_min	-0.17	0.2889	0.07	0.7108	0.15	0.4740	**-0.19**	**0.0114**	-0.11	0.1937
B2_sd	-0.09	0.5752	-0.11	0.5705	-0.18	0.3882	**-0.26**	**0.0004**	**-0.27**	**0.0010**
B3_mean	-0.19	0.2216	0.11	0.5609	0.14	0.5092	**-0.29**	**0.0001**	**-0.23**	**0.0041**
B3_max	-0.18	0.2388	-0.14	0.4832	-0.25	0.2095	**-0.33**	**9.06E-06**	**-0.28**	**0.0006**
B3_min	-0.11	0.4992	0.37	0.0537	**0.42**	**0.0319**	**-0.17**	**0.0271**	-0.09	0.2727
B3_sd	-0.14	0.3850	**-0.40**	**0.0359**	**-0.45**	**0.0205**	**-0.27**	**0.0003**	**-0.24**	**0.0030**
B4_mean	**-0.36**	**0.0175**	0.27	0.1663	**0.54**	**0.0041**	**-0.30**	**4.42E-05**	**-0.25**	**0.0019**
B4_max	**-0.36**	**0.0189**	0.06	0.7575	0.09	0.6776	**-0.38**	**2.34E-07**	**-0.35**	**1.71E-05**
B4_min	-0.21	0.1840	**0.40**	**0.0365**	**0.53**	**0.0055**	**-0.15**	**0.0433**	-0.08	0.3370
B4_sd	-0.26	0.0979	**-0.39**	**0.0381**	**-0.42**	**0.0344**	**-0.27**	**0.0002**	**-0.26**	**0.0012**
B5_mean	0.05	0.7369	0.32	0.0952	**0.44**	**0.0245**	**-0.24**	**0.0011**	**-0.17**	**0.0351**
B5_max	0.01	0.9319	0.23	0.2431	0.31	0.1187	**-0.32**	**1.50E-05**	**-0.26**	**0.0012**
B5_min	0.14	0.3578	**0.39**	**0.0409**	**0.48**	**0.0135**	-0.14	0.0678	-0.07	0.3834
B5_sd	-0.21	0.1851	-0.16	0.4140	-0.11	0.5769	**-0.24**	**0.0016**	**-0.22**	**0.0075**
Vegetation Indices										
CCCI_mean	**0.44**	**0.0031**	-0.14	0.4689	-0.29	0.1568	**0.18**	**0.0151**	**0.16**	**0.0476**
CCCI_max	**0.31**	**0.0409**	-0.26	0.1839	**-0.51**	**0.0085**	**0.19**	**0.0112**	**0.20**	**0.0132**
CCCI_min	0.30	0.0527	0.00	0.9857	-0.07	0.7201	0.07	0.3751	0.03	0.6762
CCCI_sd	-0.11	0.4736	**-0.39**	**0.0412**	**-0.51**	**0.0082**	0.13	0.0939	0.14	0.0904
NDVI_mean	0.18	0.2612	0.28	0.1421	0.35	0.0772	-0.12	0.1217	-0.05	0.5393
NDVI_max	0.12	0.4383	0.23	0.2373	0.31	0.1209	**-0.16**	**0.0323**	-0.09	0.2653
NDVI_min	0.17	0.2850	**0.40**	**0.0355**	**0.48**	**0.0126**	-0.07	0.3834	-0.01	0.9025
NDVI_sd	-0.14	0.3539	-0.20	0.3044	-0.27	0.1829	-0.09	0.2564	-0.08	0.3270
NDRE_mean	0.12	0.4420	0.29	0.1319	**0.42**	**0.0340**	**-0.19**	**0.0116**	-0.13	0.1158
NDRE_max	0.05	0.7681	0.05	0.8168	0.10	0.6441	**-0.20**	**0.0073**	-0.14	0.0889
NDRE_min	0.11	0.4703	**0.38**	**0.0477**	**0.51**	**0.0082**	-0.13	0.0942	-0.08	0.3414
NDRE_sd	-0.12	0.4592	**-0.38**	**0.0480**	**-0.46**	**0.0171**	-0.09	0.2475	-0.10	0.2355
NDVIRE_mean	**0.34**	**0.0252**	0.19	0.3370	0.25	0.2087	-0.04	0.6242	0.01	0.9065
NDVIRE_max	**0.36**	**0.0191**	0.09	0.6406	0.14	0.4845	-0.02	0.7617	0.03	0.6936
NDVIRE_min	0.25	0.1088	**0.38**	**0.0456**	0.35	0.0764	-0.03	0.6826	0.01	0.9272
NDVIRE_sd	-0.10	0.5332	-0.06	0.7457	-0.02	0.9296	0.03	0.7043	0.06	0.4721
MSAVI_mean	0.10	0.5088	0.32	0.0980	**0.41**	**0.0397**	**-0.21**	**0.0055**	-0.14	0.0966
MSAVI_max	0.05	0.7654	0.26	0.1833	0.34	0.0869	**-0.27**	**0.0002**	**-0.21**	**0.0111**
MSAVI_min	0.15	0.3241	**0.39**	**0.0375**	**0.46**	**0.0167**	-0.12	0.1102	-0.06	0.4779
MSAVI_sd	-0.19	0.2157	-0.11	0.5807	-0.05	0.8071	**-0.20**	**0.0074**	**-0.18**	**0.0264**
GLCM Texture Features										
CORR_X15	0.06	0.7193	-0.28	0.1521	-0.35	0.0788	**-0.21**	**0.0055**	**-0.31**	**0.0001**
ENTR_X15	**0.52**	**0.0004**	-0.26	0.1791	-0.38	0.0563	**-0.25**	**0.0007**	**-0.29**	**0.0003**
HOMO_X15	**-0.51**	**0.0004**	0.20	0.3036	0.35	0.0785	**0.22**	**0.0031**	**0.22**	**0.0068**
MEAN_X15	-0.22	0.1516	0.30	0.1259	**0.60**	**0.0013**	**-0.19**	**0.0117**	-0.14	0.0868

For the AWF dataset (n = 177), 27 predictors were significantly correlated with AGB, with absolute values of r from 0.15 to 0.39 ($p < 0.05$). In contrast with the EP dataset, the exclusion of coniferous-dominated plots from the AWF dataset decreased the number of predictors that significantly correlated with AGB at $p < 0.05$ and decreased the magnitude of r of the predictors (n = 148). Noticeable changes occurred in the variables from the band reflectances and vegetation indices groups. The most likely reason for the decrease in r is the reduction of AWF sample size, since calculation for the p-value of the significance level is also a function of sample size.

It was observed in this study that texture features were correlated ($p < 0.05$) more frequently with the AGB estimates of Sebangau and of Hainich with the AWF dataset, than that of Hainich with the EP dataset. The AGB estimates of Hainich with the EP dataset most frequently correlated with the band reflectance and vegetation indices predictors.

4.3.2. AGB estimation models

The final AGB model for all study cases was selected through a stepwise method with exhaustive searches and multiple linear regressions (Table 4.8). Each predictor in the model was significant for $p \leq 0.05$ and met the statistical collinearity requirements with the maximum VIF value of 5.

The variables mean CCCI and GLCM homogeneity were the two selected predictors for the Sebangau model. The model was weak, as it explained only 34% of the variability in the dataset (n = 43). The cross-validation using the LOOCV approach showed a $RMSE_r$ of 24% relative to the mean. In comparison to the models derived from the Hainich datasets, the model performance in Sebangau was slightly higher than the model derived from the AWF dataset, but lower than that of the EP dataset.

The Hainich mixed model based on the EP dataset (n = 28) consisted of three predictors. The selected predictors in the model were the maximum and minimum of the green band, as well as the standard deviation of the red-edge band. The model explained as much as 38% of the variability in the data and the $RMSE_r$ was 13.27% relative to the mean. An improvement in the model performance occurred when the coniferous-dominated plots were excluded from the EP dataset. The Hainich broad-leaved model based on the EP dataset (n = 26) comprised four combinations of predictors: the mean of the red band, the maximum of MSAVI, the minimum of CCCI, and the standard deviation of NDRE. The model explained 65% of the variability in the dataset with a $RMSE_r$ of 10.26% relative to the mean.

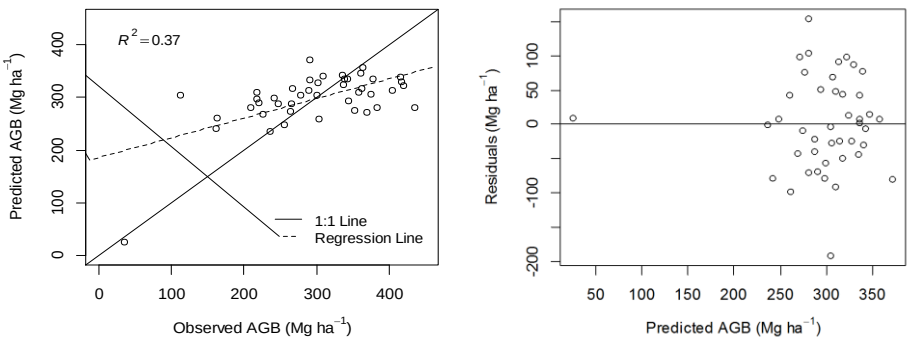
Table 4.8. Final AGB models derived from the Sebangau and Hainich datasets. The model performance is evaluated by using LOOCV; VIF = Variance Inflation Factor; $RMSE_r$ = Relative Root Mean Square Error.

AGB Model Sample size (*n*)	Predictor variables	Estimate	Std. Error	*t-value*	Pr(>|t|)	VIF	Adj R^2	*RMSEr* (%)
Sebangau model							0.34	24.01
LSI (*n* =43)	(Intercept)	-1322.3	678.6	-1.949	0.05838			
	CCCI_mean	3501.8	1332.1	2.629	0.01210	1		
	HOMO_X15	-664.9	196.0	-3.391	0.00158	1		
Hainich mixed model							0.38	13.27
EP (*n* =28)	(Intercept)	199.7	158.8	1.257	0.22070			
	B2_max	16470.3	4859.0	3.390	0.00240	2		
	B2 _min	-22222.6	8002.6	-2.777	0.01050	2		
	B4_sd	-243158.2	59784.6	-4.067	0.00040	2		
Hainich broad-leaved model							0.65	10.26
EP (*n* =26)	(Intercept)	288.0	242.5	1.187	0.24840			
	B3_mean	28659.4	11386.7	2.517	0.02000	2		
	MSAVI_max	1029.3	175.7	5.857	0.00080	2		
	CCCI _min	-1076.9	309.1	-3.484	0.00220	1		
	NDRE_sd	-10579.5	2459.0	-4.302	0.00030	1		
Hainich mixed model							0.23	42.24
AWF (*n* =177)	(Intercept)	1532.2	245.5	6.241	3.30E-09			
	B2_mean	-7058.7	2303.9	-3.064	0.00250	4		
	B4_max	-2782.6	944.4	-2.946	0.00370	5		
	MEAN_X15	1665.7	489.7	3.401	0.00080	4		
	CORR_X15	-230.3	91.5	-2.518	0.01270	1		
Hainich broad-leaved model							0.20	42.79
AWF (*n* =148)	(Intercept)	1515.7	237.8	6.374	2.30E-09			
	B2_mean	-11169.5	2704.7	-4.130	6.10E-05	3		
	B2_min	4806.8	2246.2	2.140	0.03400	3		
	CORR_X15	-364.3	105.8	-3.441	0.00080	1		

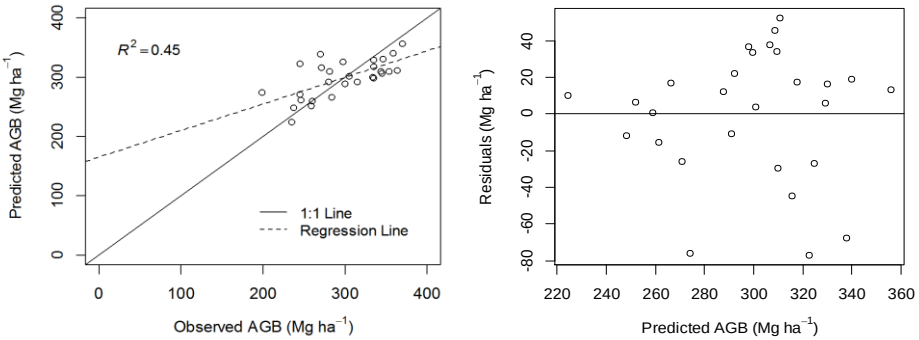
The performance of AGB models derived from the AWF dataset were lower than those derived from the EP dataset. The AWF mixed model (*n* = 177) explained only 23% of the variability in the dataset with a $RMSE_r$ of 42.24% relative to the mean. The exclusion of coniferous-dominated plots from the AWF dataset slightly decreased the model performance (*n* = 148). The decrease in the correlation between the AGB for broad-leaved and the selected predictor (B2_mean) (Table 4.7), which contributed most in estimating the AGB (Table 4.8), followed the decrease of the model performance.

Scatterplots of the observed versus predicted AGB from the Sebangau and Hainich AWF models suggest a saturation at about 300 Mg ha^{-1}. However, this was not observed for the EP model (Fig. 4.10). The highest residual was observed in the Hainich AWF models.

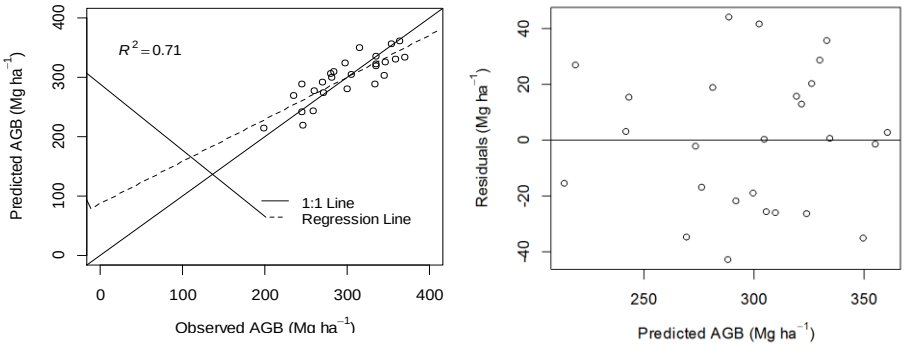
Sebangau model (*n*=43)



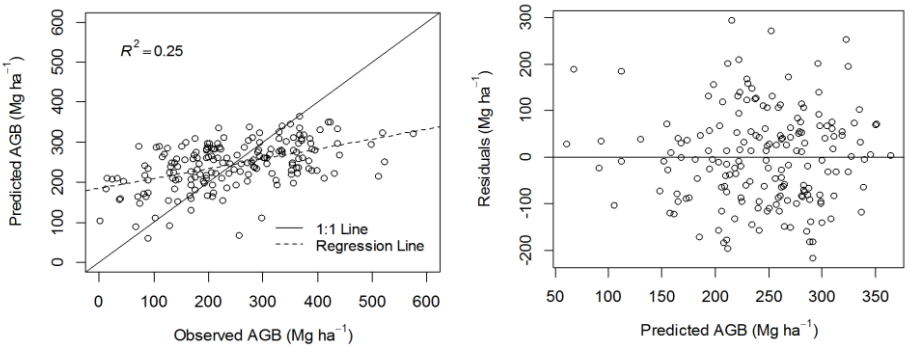
Hainich Mixed model EP (*n*=28)



Hainich Broad-leaved model EP (*n*=26)



Hainich Mixed model AWF (*n*=177)



Hainich Broad-leaved model AWF (*n*=148)

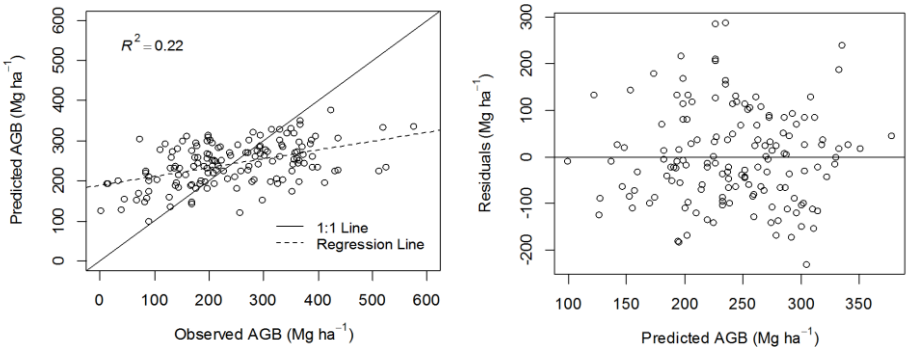


Fig. 4.10. Scatterplots of predicted versus observed values and residual versus predicted values for the Sebangau and Hainich AGB models.

4.3.3. Model regionalization and validation

Sebangau

While the model showed a weak predictive performance, the model regionalization over the Sebangau study area produced a predicted AGB map with plausible distributions, given in a 35 m spatial resolution (Fig. 4.11). The bright areas on the map represent areas with low AGB estimates, predicted to be less than 50 Mg ha^{-1}. These areas are in the forest areas the southern part of the study area, in the previously degraded/burnt areas in the western part, and in the riverine area in the northeastern part. The model-assisted mean AGB per hectare and the corresponding standard error estimate for the Sebangau study area (N = 242,735) was 281.7 ± 10.4 Mg ha^{-1}. This model-assisted estimate was within the confidence interval of the field-based mean AGB estimates reported in 4.3.1.

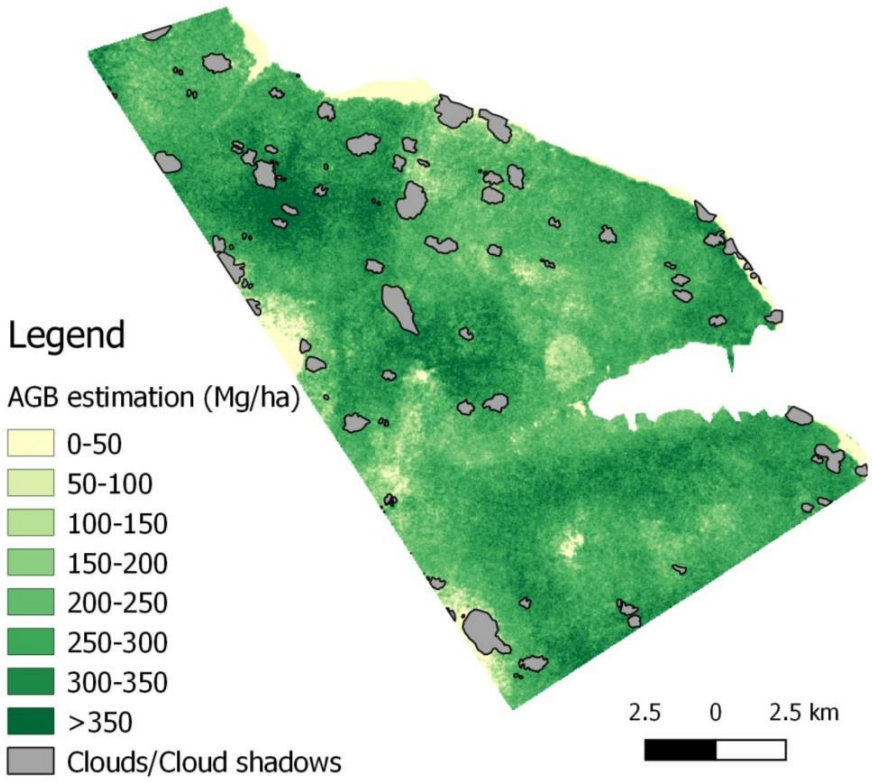


Fig. 4.11. The AGB distribution map over the Sebangau study area predicted from the model.

However, an independent validation using the HIL dataset showed that the model performed poorly at predicting AGB in the HIL plots. The Pearson correlation between the observed AGB from HIL plots and the AGB predicted by the model is just r = 0.12 (*p-value* = 0.56).

Hainich

Regionalizations of the produced AGB models from the EP and AWF datasets are presented in Fig. 4.12. Due to the smaller pixel size, the AGB-AWF map (25 m x 25 m pixel resolution) presents more detailed information of the spatial distributions of predicted AGB in the Hainich study area, compared to that of the EP-AGB map (100 m x 100 m pixel resolution). The image in the center is the RapidEye data, which was used to derive the predictor variables. Detailed information about the RapidEye image has been explained in 2.4.2. The image is presented in the false color composite (FCC) of the red-edge, NIR, and red band (FCC 543). The subset maps in Fig. 4.12 present a small part of the study area that consists of the dominant beech forests (of brownish-green appearance in the RapidEye image of FCC 543), dominant coniferous trees (very dark green), open areas (magenta), young trees (bright green), and bare lands and grass (light brown mixed with light magenta and light green). The differences in resolution between the two AGB maps are clearly visible in the subset maps.

The EP-AGB map shows higher predictions than that of the AWF-AGB map. The model-assisted mean AGB estimate for the Hainich study area based on the EP data (N = 11,780) was 265.2 ± 6.6 Mg ha^{-1}, while the model-assisted mean AGB estimate based on the AWF data (N = 188,585) was 243.4 ± 7.6 Mg ha^{-1}. This was consistent with the field-based mean AGB estimate from the EPs, which was higher than the mean estimate from the AWF plots.

Fig. 4.13. shows a map presenting the difference between predicted AGB derived from the EP and AWF datasets in the Hainich study area. The difference map was attained by subtracting the AWF-AGB map from the EP-AGB map. In the areas where the forest is dominated by coniferous trees, the AWF-AGB map generally showed higher values than in the EP-AGB map, resulting in a negative AGB prediction of up to more than 150 Mg ha^{-1}. Similar results were found in the shadowed areas of the forest, in which the AWF-AGB map predicted higher AGB than that of the EP-AGB map.

The younger stands appear as brighter green in the RapidEye image. For these areas, higher AGB values are predicted in the EP-AGB map than in the AWF-AGB map. This results in positive differences between the two maps. The EP-AGB map also showed higher AGB predictions in the open areas, e.g., open lands and grasses. A similar positive difference occurred in the forest edge along the road. Reflectances from this area were influenced by younger stands and grasses, producing higher predicted AGB of > 350 Mg ha^{-1} in the EP-AGB map. On the contrary, this area was detected with lower predicted AGB of < 50 Mg ha^{-1} in the AWF-AGB map.

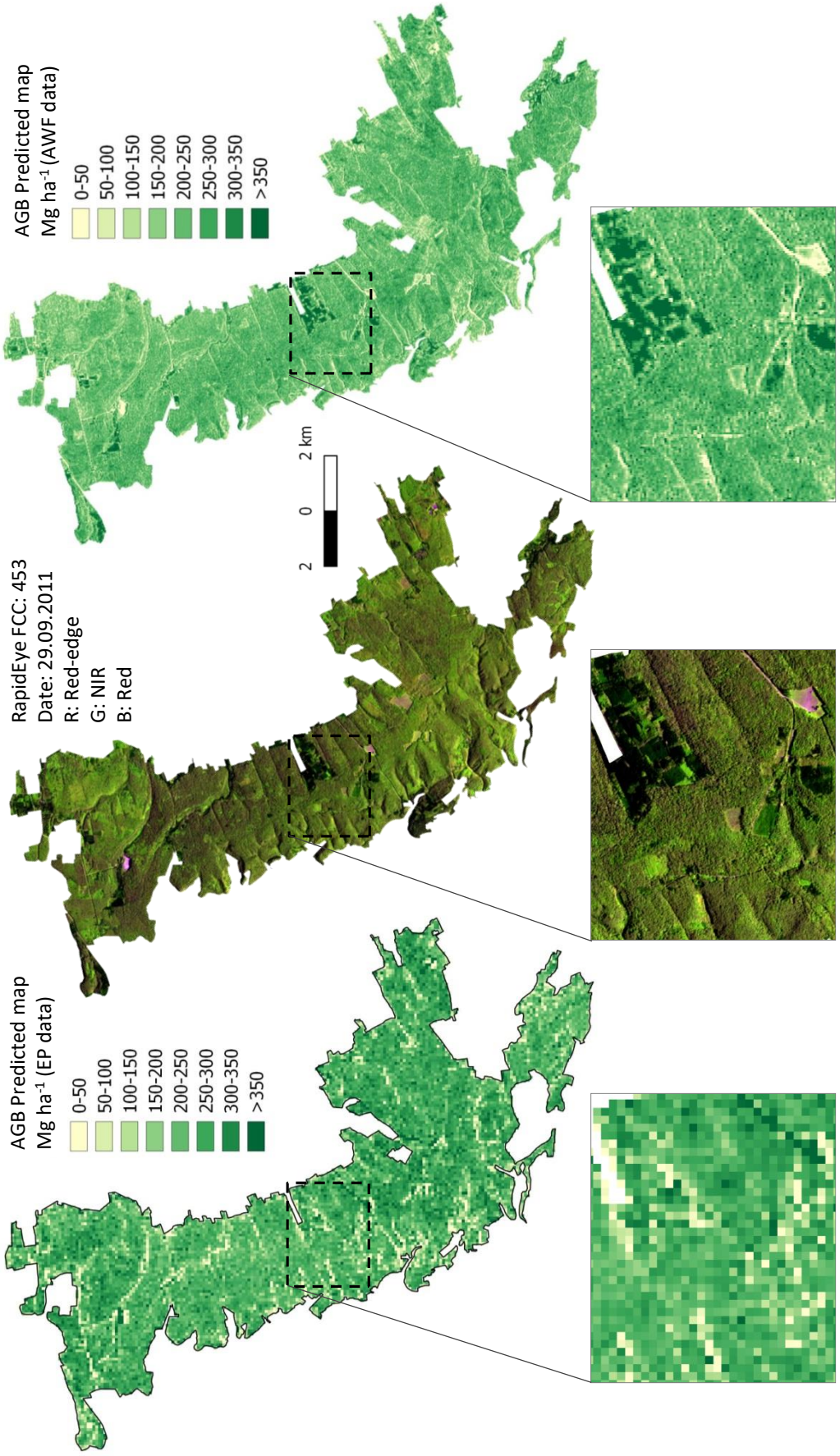


Fig. 4.12. The AGB distribution maps of the Hainich study area predicted using the model derived from the EP dataset (left), and from the AWF dataset (right). RapidEye image of the study area acquired on 20th of September 2013, presented in band combination of red-edge, NIR, and red in the red-green-blue color composite, respectively (RGB 453). Description of the subset image is given in the text.

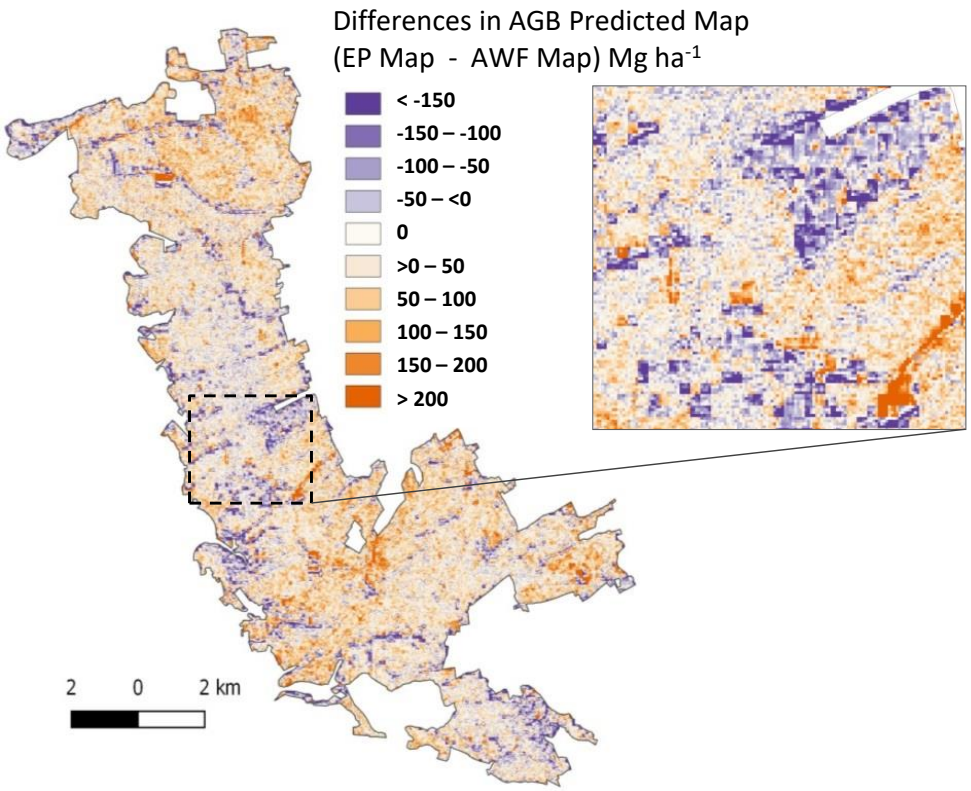


Fig. 4.13. The AGB difference map showing the difference between AGB predicted from the EP model, and the AGB predicted from the AWF model.

Results from the cross-validation of the two predicted maps show that the EP model was relatively weak in predicting the AGB of AWF plots, as well as when the AWF model was used for predicting the AGB of the EP plots. The correlation between the EP-AGB predicted map and the observed AGB in the AWF plots was $r = 0.12$ (*p-value* = 0.10), while the correlation between AWF-AGB predicted map and the observed AGB in the EP plots was $r = 0.16$ (*p-value* = 0.40), or simply no correlation.

4.3.4. Relative efficiency

Table 4.9 presents the model-assisted estimates for the mean AGB ($\hat{\mu}_{MA}$) and the associated standard error ($SE_{(\hat{\mu}_{MA})}$) and variance ($\hat{V}_{(\hat{\mu}_{MA})}$). The model-assisted mean AGB of the Sebangau and the Hainich-AWF mixed plots were within the 95% confidence interval of the corresponding estimates from the field observations. However, the mean AGB of the Hainich-EP mixed plots fell outside of the confidence interval (Table 4.6). Compared to the estimates of the *SE* of the mean from field observations, the model-assisted *SE* estimates were lower in all cases. The efficiency of AGB estimates with the integration of RapidEye data, relative to

the estimates from field data only was higher for all cases. The highest *RE* value of 3.4 was given when the RapidEye data was combined with the broad-leaved plots of Hainich-EP dataset.

Table 4.9. The model-assisted estimates for the mean AGB ($\hat{\mu}_{MA}$) and the associated standard error estimates ($SE_{(\hat{\mu}_{MA})}$) and Variance ($\hat{V}_{(\hat{\mu}_{MA})}$), and relative efficiency (*RE*).

Study area	$\hat{\mu}_{MA}$	$SE_{(\hat{\mu}_{MA})}$	$\hat{V}_{(\hat{\mu}_{MA})}$	*RE*	Pixels	
	(Mg ha^{-1})		(Mg ha^{-1})^2		Sizes (m)	*N*
Sebangau (*n* =43)	281.7	10.4	108.6	1.6	35x35	242,735
Hainich - EP						
Mixed plots (*n* =28)	265.2	6.6	43.9	1.8	100x100	11,780
Broad-leaved plots (*n* =26)	-	4.9	24.3	3.4	-	-
Hainich - AWF						
Mixed plots (*n* =177)	243.4	7.6	57.4	1.3	25x25	188,585
Broad-leaved plots (*n* =148)	-	8.3	69.2	1.3	-	-

4.4. The effect of different field plot sizes on the AGB estimation models derived from field inventory and RapidEye data using the EP dataset of Hainich

4.4.1. Estimations based on field observations of different plot sizes

Table 4.10 presents the field estimates for the mean *DBH*, stem density, estimated AGB, and estimated BA derived from the EP Dataset 1 (*n* = 26), as described in section 3.3. The mean estimation from the 95 m x 95 m and smaller plot sizes were calculated plot-wise and expanded to the 1 ha estimation using an expansion factor according to the plot size. The average number of measured trees per plot is also presented in the table for different plot sizes. The mean AGB estimates from all smaller plot sizes were within the confidence interval of the estimate from the 1 ha plot, except for the 25 m x 25 m sizes (Table 4.6). The mean stem density and *DBH* estimates increased as the plot size decreased, as well as the mean AGB estimates. As the plots get smaller, the SE of the mean AGB estimates increased, from 9.1 to 18.9 Mg ha^{-1}, while the range of AGB estimates increased from 170.9 to 402.15 Mg ha^{-1}. Likewise, the field-based SE of the mean BA estimates increased from 18.15 to 40.07 m^{2}ha^{-1}, and the range of BA estimates increased from 18.15 to 40.07 m^{2}ha^{-1}.

Table 4.10. Major forest variables from the EP datasets of different square plot sizes with the same plot center and a constant sample size of n = 26.

Plot size (m^2)	Avg trees per plot	*DBH* Range (cm)	*DBH* Mean ± SE (cm)	Stem Density Range (no ha^{-1})	Stem Density Mean ± SE (no ha^{-1})	AGB Range (Mg ha^{-1})	AGB Mean ± SE (Mg ha^{-1})	BA Range (m^2 ha^{-1})	BA Mean ± SE (m^2 ha^{-1})
100x100	386	15.72	28.7 ± 0.96	463	386.2 ± 25.11	170.94	298.6 ± 9.12	18.15	33.3 ± 1.00
95x95	353	16.07	28.7 ± 0.97	468	391.7 ± 25.32	176.65	302.9 ± 9.74	19.27	33.8 ± 1.05
90x90	316	15.81	28.9 ± 0.98	489	389.9 ± 25.95	192.24	304.9 ± 9.94	20.39	34.0 ± 1.08
85x85	281	18.80	28.9 ± 1.01	502	388.7 ± 26.03	174.11	302.0 ± 9.79	18.43	33.6 ± 1.05
80x80	249	21.68	28.9 ± 1.05	509	389.1 ± 26.40	194.32	301.6 ± 10.49	20.09	33.6 ± 1.12
75x75	219	22.77	29.0 ± 1.07	548	389.9 ± 27.57	195.53	301.4 ± 10.69	20.86	33.6 ± 1.15
70x70	191	23.54	29.1 ± 1.12	549	389.4 ± 28.26	187.75	301.9 ± 10.49	20.18	33.7 ± 1.10
65x65	165	24.47	29.3 ± 1.12	592	391.6 ± 28.97	172.77	303.0 ± 9.95	19.06	33.8 ± 1.08
60x60	141	25.63	29.5 ± 1.17	575	390.9 ± 28.94	169.04	306.9 ± 9.54	17.91	34.2 ± 1.05
55x55	117	25.99	29.8 ± 1.25	582	387.3 ± 29.99	217.35	308.6 ± 12.09	22.51	34.3 ± 1.26
50x50	95	32.80	30.3 ± 1.42	612	381.4 ± 31.46	222.27	306.4 ± 12.09	23.32	34.0 ± 1.27
45x45	77	33.91	30.7 ± 1.52	607	379.3 ± 31.70	228.66	306.2 ± 11.06	23.75	34.0 ± 1.15
40x40	60	36.49	30.3 ± 1.58	575	376.7 ± 30.74	217.12	300.0 ± 12.12	20.87	33.3 ± 1.25
35x35	46	40.24	31.0 ± 1.73	629	375.2 ± 31.74	246.92	305.2 ± 14.04	23.50	33.7 ± 1.42
30x30	34	39.65	30.9 ± 1.69	633	373.5 ± 32.48	286.03	301.7 ± 12.98	28.20	33.4 ± 1.38
25x25	24	39.74	31.3 ± 1.80	592	385.2 ± 33.08	366.55	317.8 ± 17.62	35.83	35.1 ± 1.75
20x20	16	41.52	30.9 ± 1.92	600	397.1 ± 36.49	402.15	316.8 ± 18.86	40.07	35.1 ± 1.96

4.4.2. Performance of AGB estimation models derived from different plot sizes and shapes

The performance of RapidEye-based AGB models is given in Adj R^2 and relative $RMSE_r$ (%) by using different square plot sizes derived from the EP Dataset 1 (Fig. 4.14 and Fig. 4.15) and different plot sizes and shapes from the EP Dataset 2 (Fig. 4.16). Each of the datasets had a constant sample size of n = 26 as described in section 3.3.

In Fig. 4.14, a single model with the same predictors as selected in the Hainich broad-leaved model (Table 4.8) was applied to fit a regression with varying square plot sizes. In general, the model performance declined as the plot size decreased. For example, the Adj R^2 decreased from 0.65 to 0.12 when the plot size decreased from 1 ha (100 m x 100 m) to 400 m^2 (20 m x 20 m). Some of the predictors in the model derived from smaller plot sizes were

insignificant at $p < 0.05$. The decreasing plot size increased the relative $RMSE_r$ from 10.27 to 34.04%.

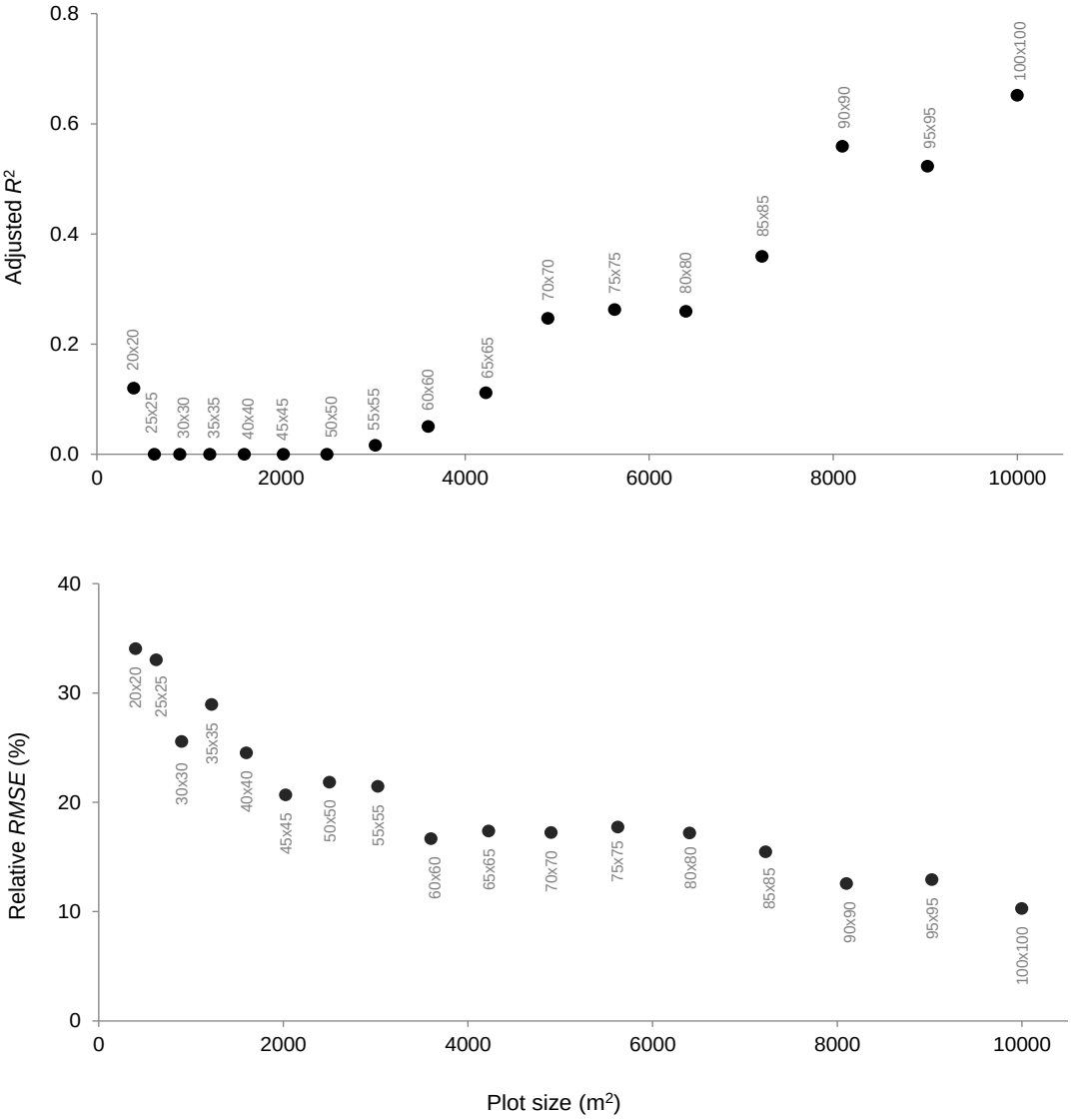


Fig. 4.14. The Adj R^2 and $RMSE_r$ (%) of the same AGB model derived from different square plot sizes with a constant center point (n = 26).

Fig. 4.15 depicts the effect of plot size in the performance of AGB models with changing predictors. The list of the separate AGB models is given in Appendix C for Dataset 1. These models explained 44 to 68% of the variability in the datasets, depending on plot size. The highest Adj R^2 of 0.68 was obtained with the 55 m x 55 m plot size. A general decreasing pattern was observed in the relative $RMSE_r$ values, while this was not the case with the Adj R^2. The lowest $RMSE_r$ was obtained from the 100 m x 100 m plot size. However, an initial steep decrease of $RMSE_r$ occurred at the 30 m x 30 m plot size with a value of 16.36% and a model

Adj R^2 of 0.54. Some results from the adjacent plots were quite similar to each other, since the square plots shared the same center point and differed only 5 m in side length.

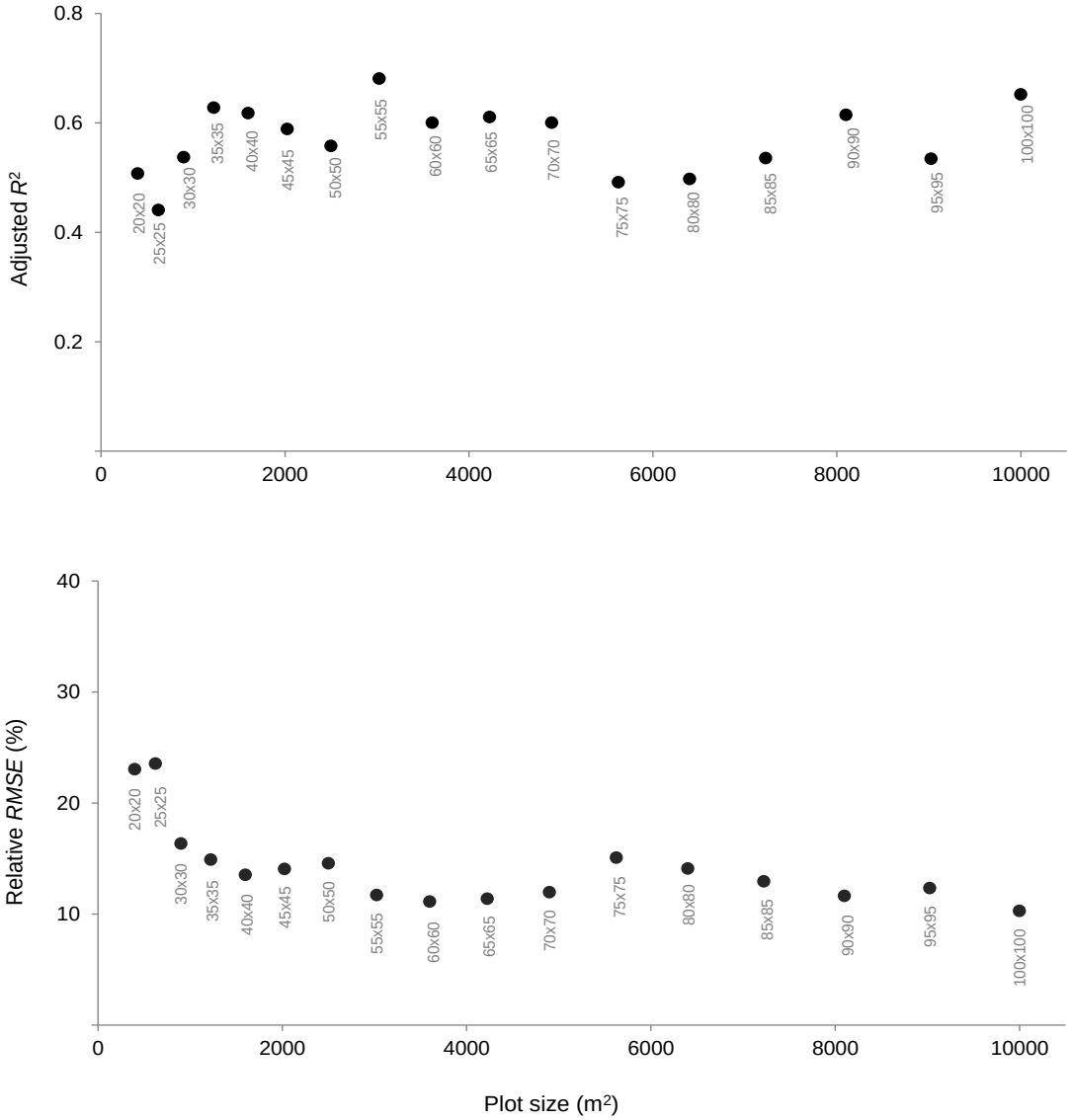


Fig. 4.15. The Adj R^2 and $RMSE_r$ (%) of different AGB models with varying predictors derived from different square plot sizes with a constant center point (n = 26).

Fig. 4.16 depicts the effects of different sample plot sizes and shapes on the performance of separate AGB models with changing predictors. The list of the AGB models is given in Appendix C for Dataset 2. Depending on the plot size, the models explained 33 to 65% of the AGB variability in the datasets. The highest Adj R^2 of 0.65 was obtained from the plot size of 100 m x 100 m. Generally, it was observed that the model Adj R^2 improved with an increase in plot size, e.g. Adj R^2 of 25 m x 20 m < Adj R^2 of 25 m x 25 m < Adj R^2 of 33.3 m x 25 m or Adj R^2 50 m x 20 m < Adj R^2 of 50 m x 25 m < Adj R^2 of 50 m x 33.3 m. However, it was also

observed that the reference plots of 20 m x 20 m, or plots with a combination side length of 20 m, e.g., 33.3 m x 20 m, 50 m x 20 m, tended to have lower Adj R^2 compared to other plots of comparable sizes. Hence, no clear direction between Adj R^2 and plot size can be drawn in this study.

Regarding the relative $RMSE_r$ values, a general decreasing pattern was also observed in Fig. 4.16 with the increase in plot size. The lowest $RMSE_r$ of 10.27% was obtained from the plot size of 100 m x 100 m. However, the initial steep decreasing rate of the $RMSE_r$ occurred when the plot size was 25 m x 25 m with the value of 20.76% and the model Adj R^2 of 0.53.

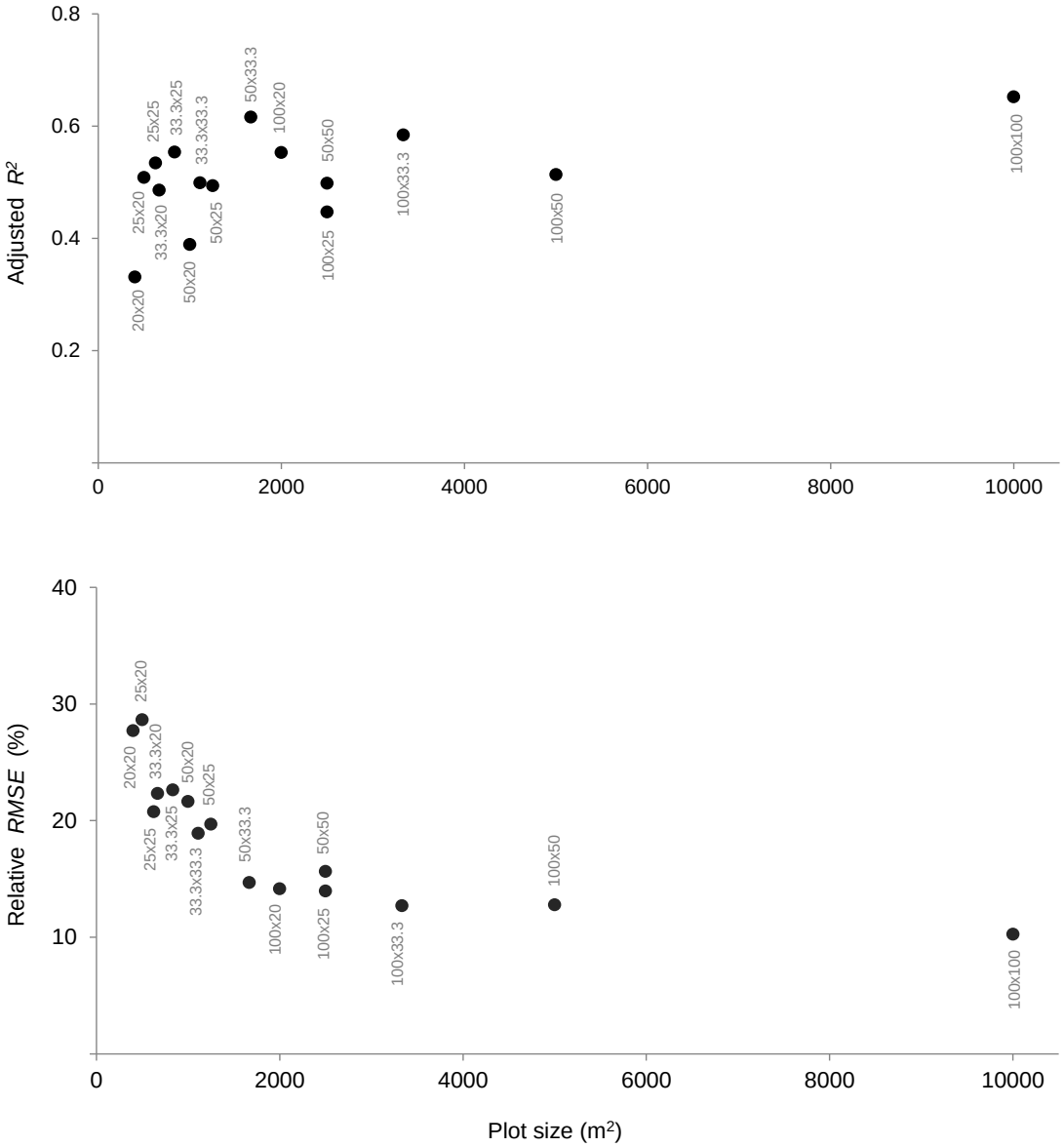


Fig. 4.16. The Adj R^2 and $RMSE_r$ (%) of different AGB models with varying predictors derived from different plot sizes and shapes selected randomly within the 1 ha EP plots (n = 26).

Statistically significant relationships were identified when correlating the AGB model performance with plot perimeter and plot size, using Datasets 1 and 2 and changing predictors. The Pearson's correlation coefficients between perimeter and relative $RMSE_r$ were higher than those between plot size and relative $RMSE_r$. In addition, the same results were identified with relative maximum residual (Table 4.11). The least correlation was found between both plot perimeter and plot size when paired with the model Adj R^2. However, these correlations had higher significance in Dataset 2 with more variations on plot shape, compared to Dataset 1 with only the square plot shape. Altogether, the results in this study indicate that plot perimeter has a higher effect in the model performance than plot size.

Table 4.11. Pearson correlation coefficients between plot perimeter (or plot size) and model Adj R^2, Relative $RMSE_r$, or Relative maximum residual derived from different plot sizes (n = 26) using Dataset 1 and Dataset 2 of 17 and 15 different plot size, respectively.

	Plot perimeter		Plot size	
	Correlation	*p-value*	Correlation	*p-value*
Based on the EP Dataset 1				
Adj R^2	0.18	0.47950	0.13	0.62710
Relative $RMSE_r$	-0.70	0.00160	-0.61	0.00937
Relative Maximum Residual	-0.82	5.37E-05	-0.78	0.00025
Based on the EP Dataset 2				
Adj R^2	0.53	0.04135	0.55	0.03521
Relative $RMSE_r$	-0.89	1.12E-05	-0.72	0.00225
Relative Maximum Residual	-0.84	8.73E-05	-0.74	0.00149

4.4.3. Efficiency of AGB estimates with the integration of RapidEye data, relative to the estimates from field data alone, for different plot sizes

Fig. 4.17 depicts the estimated SEs of the field-based mean AGB estimates and the model-assisted mean AGB estimates, obtained from Dataset 1. The estimated SE from the model-assisted estimates was lower than that of the field-based estimates. Both estimations showed a decreasing trend with increasing plot size, with the lowest estimated SE from the 100 m x 100 m plot size. However, the initial steep decreasing rate of the estimated SE from both approaches occurred at a plot size of 30 m x 30 m.

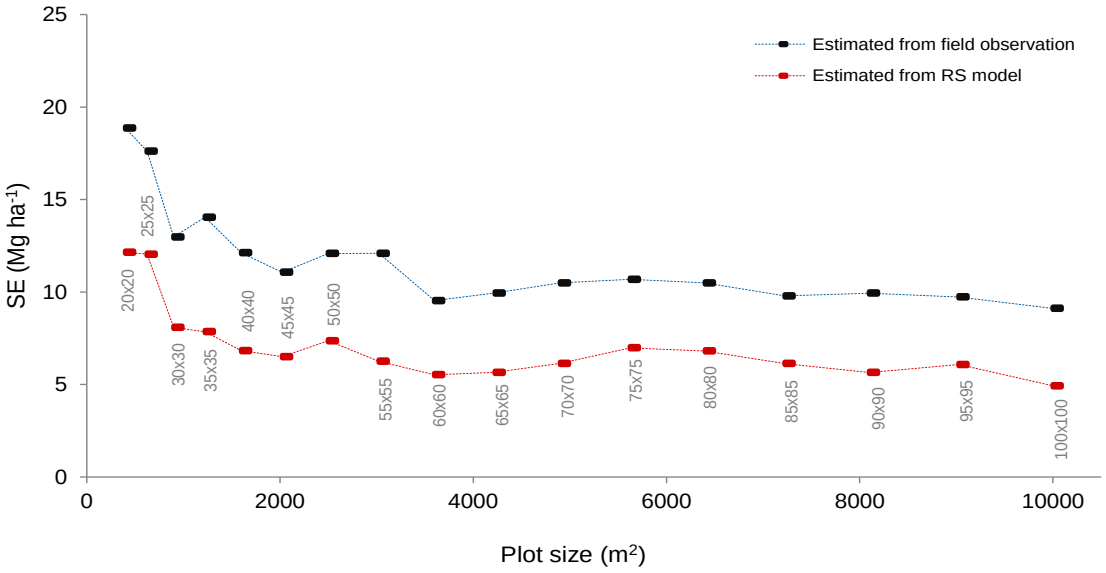


Fig. 4.17. The standard error (SE, Mg ha^{-1}) of different AGB models at different square plot sizes with the same center point.

Fig. 4.18 depicts the relative efficiencies of the two approaches. The values range between 2.1 and 3.7, with an average value of 2.8. The result showed that the precision of the mean AGB estimates improved when the RapidEye data was incorporated with the field observation data, compared to the precision from the field-based mean AGB estimate alone. However, the magnitude of this efficiency was different for different plot sizes with a similar oscillation pattern shown by the magnitude of the model Adj R^2.

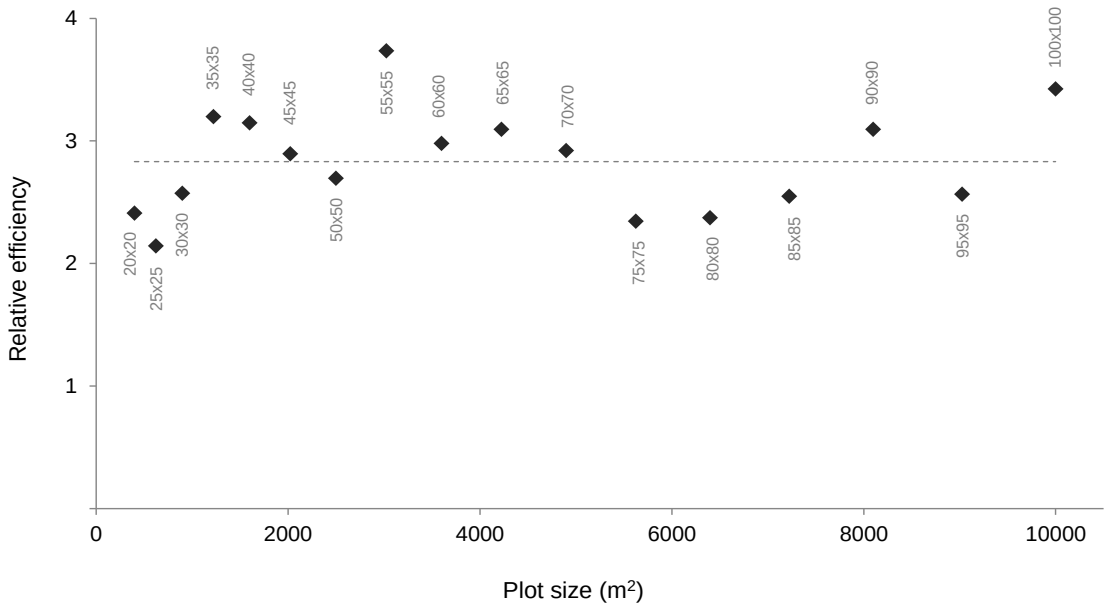


Fig. 4.18. The relative efficiency (unitless) of different AGB models from different square plot sizes with the same center point. The line shows the average efficiency value.

The results from the Dataset 1 indicate an optimum plot size of 55 m x 55 m. The model derived from this plot size exhibits a high Adj R^2, low $RMSE_r$, low SE of the mean AGB estimate, and high efficiency from the inclusion of RapidEye data in the estimation. However, since a smaller plot size is usually preferable for field data collection, the 40 m x 40 m plot size can be taken as an alternative for the optimum plot size, depending on the available sources and the acceptable precision.

Comparable results were achieved when the analysis was carried out using the Dataset 2, as it was depicted in Fig. 4.19. The estimated SE of the field-based mean AGB estimates ranged between 9.1 to 22.8 Mg ha^{-1}, and the estimated SE of the model-assisted mean AGB estimate was within the range of 4.9 to 14.6 Mg ha^{-1}. While the initial steep decreasing rate of the estimated SE from both approaches occurred when the plot size was 25 m x 25 m, the lowest estimated SE was achieved from the plot size of 100 m x 100 m.

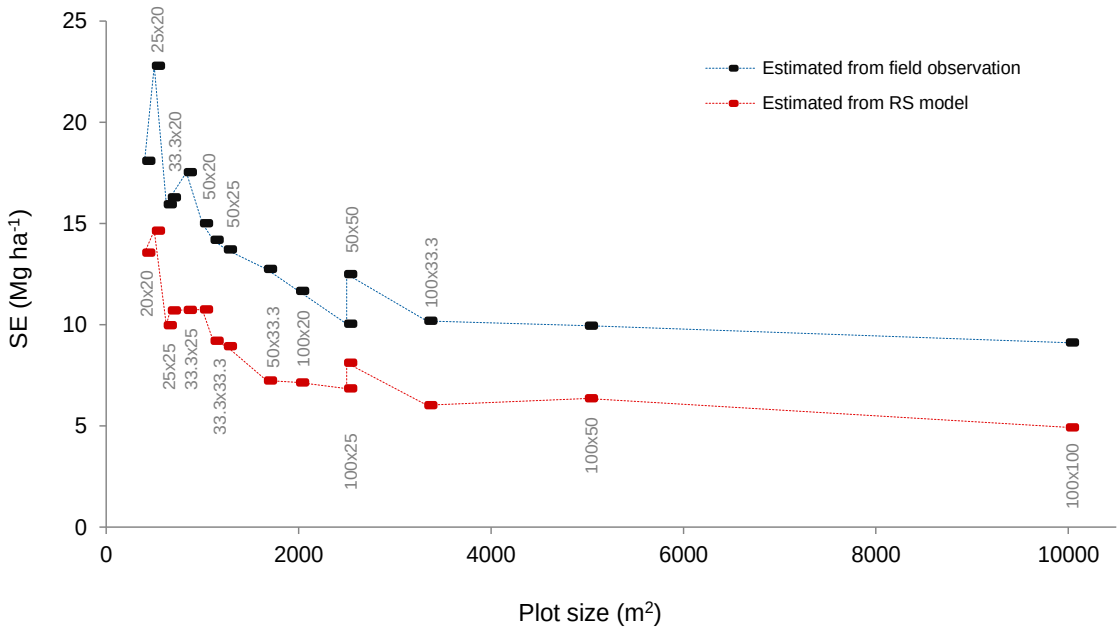


Fig. 4.19. The standard error (SE, Mg ha^{-1}) of different AGB models from different plot sizes and shapes, selected randomly within the 1 ha EP plots.

Fig. 4.20 depicts the relative efficiencies between the two approaches. The values were between 1.8 and 3.4 for different plot sizes, with an average value of 2.5. The highest relative efficiency of 3.4 was obtained from the plot size of 100 m x 100 m. A similar oscillation pattern as the one shown by the Adj R^2 was also found in the relative efficiency of the two approaches.

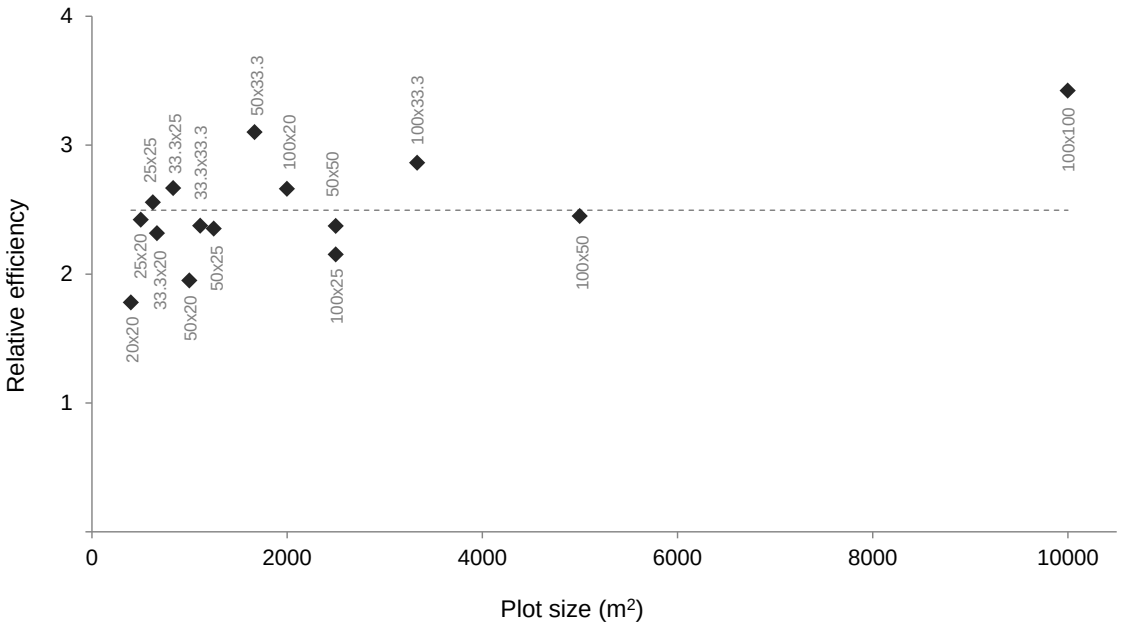


Fig. 4.20. The relative efficiency (unitless) of different AGB models from different plot sizes and shapes, selected randomly within the 1 ha EP plots. The line shows the average efficiency value.

From the results using the Dataset 2, the optimum plot size with a combination of high Adj R^2, low $RMSE_r$, low SE, and high efficiency from the inclusions of RapidEye data in the estimation, was the 50 m x 33.3 m plot size. Considering the available sources and the acceptable precision, a plot size of 25 m x 25 m can be taken as an alternative.

Overall, the results showed that the AGB model with a higher Adj R^2 exhibits a larger gain in the precision of AGB estimates by combining RapidEye and field data.

Chapter 5

Discussion

5.1. The effect of forest restricted visibility on the basal area estimates from ACP in two different forest types of Sebangau and Ncaute

FAP and ACP designs are commonly used in forest inventories for the estimation of forest stand variables. In the absence of measurement error, both approaches produce unbiased estimates of, for example, basal area. When applying the ACP method, the trees missed to be counted at a larger distance are often the most relevant sources of measurement error, depending on the selected BAF and the visibility conditions of the forest during the measurement.

The maximum distance of visibility of the Sebangau forest in this study was 6.6 m. Given this restricted visibility, a BAF of 5 was used per the standard in dense forest conditions like in Sebangau. This BAF was low enough to be largely unaffected by the restricted forest visibility and high enough to provide a high precision of estimation. Results from the simulation study indicated a systematic underestimation of the basal area from ACP in the Sebangau forest, particularly when a BAF < 5 was selected. These results were interpreted as the effect from the restricted visibility within the forest. The use of a lower BAF required a greater distance of visibility than the maximum distance of visibility of the forest, which led to a deviation of the estimations. The results were in accordance with earlier studies (Brooks, 2006; Husch et al., 1982; Ritter et al., 2013). The study of Ritter et al. (2013) found that basal area estimates from ACP with a BAF of 1 or 2 were biased, due to the increased number of miscounted trees at a larger distance. Their study was conducted by using a simulated forest with trees spatially distributed as random and Poisson point patterns. On the contrary, selecting a BAF > 5 will inflate the rate of basal area estimates represented by each sample tree. The use of higher BAFs reduces the measurement error due to the missing counted trees, but it also tends to increase the sampling error of the estimations (Zeide, 1985).

This study estimated mean basal area from FAP of 31.0 m^2ha^{-1} in the Sebangau forest and of 3.6 m^2ha^{-1} in the Ncaute forest. Kronseder et al. (2012) reported a similar mean basal area for the Sebangau forest of (mean ± SD) 30.7 ± 12.4 m^2ha^{-1}, obtained using the ACP approach with

no information on the BAF used in the study. The sample plots of Kronseder et al. (2012) covered a wider range of AGB variabilities from unlogged to slightly-logged forests. This can be the reason for the higher standard deviation compared to the one in the current study of 8.4 m^2ha^{-1}. No further comparison can be made since the study did not report the standard error of the estimation.

The difference between the estimates from ACP and FAP in the field inventory data also indicated the effect of forest restricted visibility in basal area estimation from ACP in the Sebangau forest. As the BAF increased, the difference between the mean basal area estimates from both approaches decreased. This interpretation was motivated by the contrasting result from Ncaute open savanna forest with hardly any visibility restriction. In line with earlier studies (Piqué et al., 2011; Scrinzi et al., 2015), the estimates from Ncaute forest showed no significant deviation between the mean estimate from ACP and FAP. These studies were conducted in managed forests with much lower tree density and less undergrowth than in the Sebangau forest, where a relatively longer distance of visibility is expected. Piqué et al. (2011) found no significant differences between FAP and ACP in the mean estimates of the basal area of different Mediterranean forest stands of Catalonia, north-east Spain. The study of Scrinzi et al. (2015) found a correlation of $r = 0.92$ between the basal area estimates from ACP and the estimates from FAP. The study was conducted in the Latemar forest of Bolzano Province in the Italian Central Alps.

Among the important criteria for ACP is the average number of trees per sample point. In the Sebangau field inventory data, the use of BAF 4 provided, on average, between 5 to 16 counted trees per sample point, as suggested by earlier studies (Avery and Burkhart, 2002; Bitterlich, 1984; Van Laar and Akça, 2007; West, 2009). In contrast, BAFs from 1 to 4 did not provide a sufficient number of trees per angle count-plot in the Ncaute forest. An average number of less than 4 trees were obtained from all three BAFs in the Ncaute forest. This result indicated that a BAF of 1 was already too high for the forest.

In the end, selection of a proper BAF is a trade-off between minimizing the deviation of the mean estimates and achieving the highest possible precision. The result of using a BAF of 5 in this study is different than the current practice applied in the temporary plots of the Indonesian NFI, in which a BAF of 4 is selected independent of the forest conditions (Indonesia MoF, 1992). The objective of the measurements on the NFI temporary plots is to obtain information on forest stock conditions (i.e. wood supply) at the observation point in time. There was no further information available on the selection procedure of this BAF, except that the target trees to be counted had a minimum *DBH* of 20 cm. However, by using the dataset and

following the approach of the current study, selecting the BAF 4 for a logged-over secondary peat-swamp forest like the Sebangau would significantly underestimate the basal area of the forest, as it was found here in the field inventory data and the results from the simulation study.

Regarding the Ncaute forest, this study found that the use of a BAF lower than 1 would be suitable for a forest type like the Ncaute. Lowering the BAF down to 0.25 m^2ha^{-1} in an open savanna forest with a low tree density and large *DBH* had also been recommended by Bitterlich (1984).

This study successfully determined the maximum visibility distance of the Sebangau forest and the suitable BAF to be used for implementing the ACP approach. The tree map used for the simulation study was based on the measured trees of a $DBH \geq 5$ cm. In practice, trees with a $DBH < 5$ cm can also be selected in the field when implementing the ACP method. However, the chance of these trees to be selected during the field survey was assumed to be so low that it could be neglected. Since trees with a $DBH < 5$ cm were not tagged in the plot, they were recognized as the non-target trees of the study.

The simulation assumed an equal visibility distance in all directions. However, this assumption may not apply in the field since natural forests stands are not necessarily isotropic. The assumption holds in the field because observers may step aside to get a better visibility for counting the sample trees.

5.2. Integration between field inventory and optical RapidEye data for the AGB estimation models in two different forest types of Sebangau and Hainich

5.2.1. Relationships between field AGB estimates and predictor variables derived from RapidEye data

The objective of the current study was to develop an AGB model derived from field inventory and RapidEye data using the Sebangau forest as the case study. The model was compared to the AGB models derived for the Hainich forest using two different field datasets. Results from this study were also compared with the results from earlier studies conducted for the development of a forest AGB estimation model. Some common statistical values such as the mean and its corresponding standard error or standard deviation (when the standard error was not available), the model coefficient of determination R^2, and the model prediction accuracy of *RMSE* have been used as performance indicators across various studies. However, these

studies may have used different remote sensing sensors, as well as sampling and statistical approaches to achieve their model.

A number of studies on the AGB model development using field inventory and various remote sensing data have also been conducted in the Sebangau forest area (Ballhorn et al., 2011; Englhart et al., 2011, 2012b, 2013; Jubanski et al., 2013; Kronseder et al., 2012; Magdon et al., 2018; Schlund et al., 2015). Except for the study of Magdon et al. (2018), none of the studies were conducted using a probability sampling design, as it was implemented in the current study. The probability sampling design allows statistically sound estimates of the target population parameters and the corresponding standard error (Köhl et al., 2006).

This study found that the field-based mean AGB estimate of the Sebangau peat-swamp forest was 297.7 Mg ha^{-1}. An earlier study by Kronseder et al. (2012), which was conducted in the same forest, reported a lower mean AGB estimate of 258.5 Mg ha^{-1}. Different results between the two studies can be explained by the wider range of AGB values in Kronseder et al. (2012), consisting of plots that were sampled from a larger area than the current study. The use of different allometric models is reported as one of the most important sources of uncertainty in estimating the forest AGB (Picard et al., 2015; Sileshi, 2014). This feature also contributed to the different AGB estimates between the two studies. Kronseder et al. (2012) used the general allometric model from Chave et al. (2005) to calculate the AGB estimates, whereas a specific allometric model for tropical peat-swamp forests was applied in the current study. The use of general allometric models is suitable for local AGB estimation in tropical remote sensing studies (Næsset et al., 2016), primarily when the local models do not represent all the sample trees, e.g., large trees. Since the local allometric model employed in the current study covered the ranges of *DBH*, height, and wood density of all trees in the observed sample plots, it was considered more appropriate.

A recent study of Magdon et al. (2018) reported a field AGB estimation of the Sebangau forest of 241.4 Mg ha^{-1} and a standard error of 4.6% relative to the mean. The study covered a small portion of the study area with easy access to the forest. This explained the lower AGB estimates of their study compared to the estimate in the current study.

The correlations between the AGB estimates and the predictor variables from the RapidEye bands were mostly negative in both the Sebangau and Hainich study areas. These inverse relationships are explained by the large proportions of incident solar radiation in the visible spectral region, between 400 to 700 nm, which is absorbed by the pigments at canopy level due to the photosynthesis process (Jones and Vaughan, 2010).

Some of the reasons for the weak relationships between the field-based AGB estimates and predictor variables in the Sebangau forest are related to the high stand density and high diversity of the tree species, in which AGB variations cannot be detected by RapidEye data. In the case of the Hainich forest, the high heterogeneity of the forest stand structure was a contributing factor. A relatively homogeneous absorption level of the red photosynthetic spectrum was observed in the Sebangau dataset. Earlier studies that were conducted in different forest types reported similar results and determined that dense and complex forest stand structures led to the saturation of the reflectance from optical data. This is one major drawback in estimating AGB (Lu et al., 2012, 2014; Pfeifer et al., 2016; Wijaya et al., 2010b).

The red-edge band shows significant correlation with the AGB variability, particularly in the Sebangau dataset. The importance of the RapidEye red-edge band for AGB estimation using remote sensing data was also reported by Næsset et al. (2016) and Pfeifer et al. (2016). In general, GLCM texture features possessed higher correlations with the AGB of all datasets. The results confirmed reports from earlier studies on the importance of biomass predictor variables derived from the spatial variability of neighboring pixels in dense and complex forest structures (Castillo-Santiago et al., 2010; Eckert, 2012; Lu et al., 2014; Sarker and Nichol, 2011; Wijaya et al., 2010a). When the optical remote sensing data mainly captured the information at canopy surface including the associated canopy shadows, the GLCM texture features reduced the shadow effects at the canopy layers (Lu, 2005). These shadow effects are attributed to the heterogeneity of stand structures, tree species composition, and canopy gaps. Low spatial resolution data has less potential for textures (Wulder, 1998). Hence, the relatively high spatial resolution of RapidEye is considered as one of the advantages to derive the GLCM texture features from the data. The study of Næsset et al. (2016) also confirmed the ability of RapidEye data to capture spatial heterogeneity from the pixels.

The inverse correlations shown with the red-edge and NIR reflectance in the EP and AWF datasets are explained by the young forest stands of the AWF plots. The more vigorous growth and denser canopy of young forest stands produces higher reflectance in the red-edge and NIR than the reflectance from the canopy of an old-growth stand (Jones and Vaughan, 2010). In old-growth stands, large trees generate more shadows on the canopy, contributing to the decreased relationships between forest stand parameters and remote sensing spectral responses (Lu et al., 2004). The increase in NIR reflectance followed the decrease in AGB estimates from the AWF plots but showed the opposite trend for the EPs. Therefore, the exclusion of coniferous-dominated plots slightly decreased the correlation between some predictor variables and the AWF-AGB estimates. In contrast, the exclusion of two coniferous-

dominated plots from the EPs produced higher correlation between predictor variables and the AGB estimates. The reduced sample size of AWF plots was the most likely reason for the decrease in the correlation, since the *p*-value calculation is a function of sample size.

5.2.2. AGB estimation models from field and RapidEye data

As the relationships between the field-based AGB estimates and the predictor variables were weak, the model performance derived from the two datasets in the current study was also poor. By using the data and following the procedure applied in this study, the resulting AGB model in the Sebangau forest explained about 34% of the AGB variability in the dataset with a relative $RMSE_r$ of 24%. This coefficient of determination was much lower than that obtained by Englhart et al. (2012a), with R^2 of 0.92, which was conducted in the same forest also with RapidEye data. Another study from Næsset et al. (2016) also reported a higher Adj R^2 of 0.53 of the AGB model obtained for the Miombo woodland, Tanzania, by using RapidEye data. The higher R^2 of the models from Englhart et al. (2012a) and Næsset et al. (2016) could be explained by the wider range of AGB variability in their field reference plots. While the present study focused on the logged-over peat-swamp forest, the reference plots of Englhart et al. (2012a) spanned AGB variation of heavily degraded and regrown areas to peat-swamp forests, and the reference plots of Næsset et al. (2016) covered the AGB variation of open land, grassland, and woodland to forest. This wider range of AGB variation within the field dataset could not be captured completely by the variability of the remote sensing signals and led to a reduction of the estimation accuracy. This is one reason for the higher $RMSE_r$ of the model from Englhart et al. (2012a) of 44%, and from Næsset et al. (2016) of 62.7%, compared to the relatively more homogenous stand structures in the Sebangau forest. Different predictor variables and different statistical approaches for producing the models were also among the possible reasons of the different model performances among the three studies.

Given a large amount of unexplained AGB variability in the Sebangau dataset, the predicted AGB values from the model regionalization are far from the field-based AGB estimations of the HIL plots. For example, while observing the model residuals, the plot with the largest negative residual contained many large trees and the model underestimated the AGB of this plot. On the other hand, in the plot with the largest positive residual, many small and medium trees were observed.

In the case of Hainich study, the AGB models for the mixed and broad-leaved forests derived from the EPs were superior to those derived from the AWF plots, particularly with respect to

the *RMSE* values. These results are attributed to the larger EP plot sizes. Earlier studies have reported that larger plots are likely to capture more within-plot variability and reduce the variance among plots, thereby increasing the prediction accuracy of the model (Köhl et al., 2006; Mascaro et al., 2011; Næsset et al., 2013).

For the Hainich mixed AGB models, 77% of the AGB variability in the AWF dataset and 62% in the EP dataset remained unexplained. The two maps predicted different levels of AGB in the Hainich forest with almost no correlation between the predicted AGB on the EP-AGB map and the observed AGB from AWF plots. Also, no correlation was observed between the predicted AGB on the AWF-AGB map and the estimated AGB from EPs. Regarding spatial resolution, the predicted AWF-AGB map presents more detailed and precise information. The smaller pixel size of the predicted AWF-AGB map was expected to show greater variations of AGB in the study area compared to that of the EP-AGB predicted map.

The smaller range between the minimum and maximum values of AGB in the EPs compared to the AWF plots limited the predictive ability of the EP-AGB model. The EP-AGB model was likely to under- and overestimate the AGB that fell outside of the model range. On the other hand, the wider range of AGB variability in the AWF dataset also influenced the accuracy of AGB predicted from the AWF-AGB model due to the saturation problem of the RapidEye data.

The AWF-AGB map predicted high AGB values upwards of 350 Mg ha^{-1} for the areas dominated by coniferous trees and shaded areas. This result was attributed to the reflectance from the AWF plots that received higher influence due to shadow effects. The study from Hansen et al. (2015) reported that the shadow effect is higher in stands with larger tree crowns and greater in smaller plots. Accordingly, some of the broad-leaved-dominated plots in the AWF plots produced similar reflectances with the coniferous-dominated plots. This argument was supported by the overestimation of AGB values in the sample plots dominated by coniferous trees. This overestimation decreased after the exclusion of the coniferous plots. Hence, the exclusion of coniferous-dominated plots slightly reduced the model performance derived from the AWF plots, while the opposite was true for the model derived from the EPs.

The AWF broad-leaved model derived from the AWF plots explained only 20% of the AGB variability in the dataset. On the contrary, the broad-leaved AGB model derived from the EPs explained 65% of the AGB variability in the dataset. Moreover, more variability was explained in the EP broad-leaved model than in the mixed forest model. This result is in accordance with the results from Latifi et al. (2012) and Wallner et al. (2015) that stratification based on forest types generally improved the estimation of forest structures and forest AGB.

The use of the EP and AWF datasets produced different descriptions of the Hainich forest with different mean AGB estimates of the forest, reaching up to 50 Mg ha^{-1}. The EPs produced a lower standard error of mean AGB than the AWF plots. This result is explained by the larger size of the EP plots, which ultimately reduced the variance among plots. Based on the field inventory datasets, the EPs were likely representing a more homogenous stand structure than those of the AWF plots. The AWF dataset showed a higher standard deviation than the EP dataset due to the higher AGB variability in the AWF dataset.

Different results from the EP and AWF datasets also arise from the differences in plots design. The concentric sub-plot design of the AWF dataset is known as an efficient approach for selecting sampled trees to be measured in forest inventory. In the development of an AGB model using a concentric sub-plot design, remote sensing information is usually extracted according to the size of the largest sub-plot and linked with the observed AGB of the plot. This observed AGB is estimated from subsample trees that are measured in each sub-plot according to their *DBH* threshold and then transformed into corresponding unit areas. The uncomplete measurement of trees in the circular plot may affect the performance of the AGB model developed from the field sampling and remote sensing data (Næsset et al., 2016).

The AWF Hainich demonstrated a higher $RMSE_r$ (42.2%) than the Sebangau dataset (24.4%). The result could be attributed to the higher AGB variability of the AWF dataset arising from the higher heterogeneity of the stand structures among plots. According to the AWF dataset, 19%, 41%, and 40% of the forest basal area were derived from trees within the *DBH* range of 7-25 cm, 25-50 cm, and >50 cm, respectively. For the Sebangau dataset, 73%, 24%, and 3% of the forest basal area were derived from trees within the *DBH* range of 7-25 cm, 25-50 cm, and >50 cm, respectively. Referring to these stand structures, a higher influence from the shadow effects on the canopy level was expected in the AWF plots than in the Sebangau plots. However, future studies on the influence of stand structure on the shadow effect at canopy level are needed.

It was also observed that 16% of the AWF dataset contained ≤ 7 sample trees per plot, which was not the case in the Sebangau plots. Two of the AWF plots contained only one single tree, with *DBH* values of 8.6 cm and 90.2 cm in each plot. This low number of sample trees led to a higher variance across sample plots and yielded a lower precision of estimation. Moreover, the influence of soil background on plots with low numbers of trees can significantly contribute to the total reflectance of the plots. This effect may lead to a wider range from the reflectance of all plots and reduce the correlation between the plots and the remote sensing variables. All

aforementioned reasons may contribute to the lower $RMSE_r$ of the AGB model derived from the AWF dataset when compared with the one derived from the Sebangau dataset.

The results demonstrate improved performance of AGB estimation with larger plot sizes, consistent with studies conducted using different sensors, such as microwave radar (Saatchi et al., 2011), LiDAR (Gobakken and Næsset, 2009; Mascaro et al., 2011; Mauya et al., 2015; Zolkos et al., 2013), or the combination of both (Næsset et al., 2015). In general, larger field plots will cover larger overlapping areas between the field and remote sensing data, which will compensate for co-registration problems (Gobakken and Næsset, 2009). An increased plot size also has a tendency to reduce model prediction error, due to the spatial averaging of the error (Tomppo et al., 2017). However, larger field plot sizes always means higher costs per plot. Thus, selection of a suitable plot size is always a compromise between the field efficiency, the expected precision within the defined efficiency level, and the available resources.

This study confirms the limitation of optical remote sensing data for extracting the AGB information in complex forest stand structures due to the insensitivity of the data to the change of AGB variations below the canopy layers. The largest negative residuals depict an underestimation in plots consisting of very large trees. In contrast, the plots with the largest positive residual contained many small trees, and the model overestimated the AGB. However, the results also imply that the saturation problems tend to have lower effects in larger plots like the EPs.

5.2.3. Relative efficiency

The inclusion of RapidEye data, in addition to field data, improved the precision of AGB estimates, indicated by the *RE* values of > 1 for all AGB models. The highest improvement occurred in the 1 ha Hainich-EP broad-leaved plots. An earlier study of Mauya et al. (2015) also reported that the gain in *RE* is higher as plot size increases.

This study shows that the gain in *RE* depends on the relationships between the RapidEye predictors and the field reference AGB estimates. This explained the similar gain in *RE* when the RapidEye data was combined with the Hainich-AWF mixed model or with Hainich-AWF broad-leaved model. When remotely sensed data is employed to assist in the estimation of AGB, a good model of fit is important to improve the gain in precision of the estimation.

5.2.4. Other common source of errors

Other sources of error may have influenced the AGB model performance. Potential errors in the field datasets include inaccuracy of the field measurements from human errors, measurement devices error, and discrepancies from different surveyors. High forest stand density made height measurements more challenging. The crowns of neighboring trees usually overlapped one another, and sometimes it was difficult to see the individual treetops. Measurement errors from *DBH*, tree height, and tree wood density due to misidentification of the tree species were also reported by Molto et al. (2013). To minimize such impacts, on-the-job training of the measurement procedures and the standard use of the measurement devices was conducted for the surveyors in the Sebangau forest prior to the field survey. However, these errors cannot be removed entirely, with some remaining at an unknown level.

The integration of field and optical remote sensing data was a complex process. There are factors influencing the interaction of radiation with leaves at the canopy level, including the incident angle of radiation with respect to the leaf and the leaf arrangement in the canopy (Jones and Vaughan, 2010). Variations of the reflectance in an image may also not represent the variations of the object reflectivity. This discrepancy can be due to local differences in irradiance over the entire image, e.g., areas with shadows (Jones and Vaughan, 2010).

One of the major challenges with optical remote sensing data is capturing a clear image that is free of clouds and other atmospheric conditions that obscure the field sample plots. Therefore, a time gap between image acquisition date and the period of field data collection commonly takes place despite the potential error that may be introduced. The difference between image acquisition date and the field data collection was within 9 months in the Sebangau study, while it was within 2 and 4 years for the AWF and EP datasets for the Hainich study. The RapidEye tiles of the Hainich study were acquired from September 29, 2011, as it was latest data available for this study. During this period, some leaves from deciduous trees might start changing color. Since all the tiles had the same date of acquisition, similar effect from the seasonal environmental conditions can be expected for the entire study area.

5.3. The effect of field plot sizes on the AGB estimation model derived from field inventory and RapidEye data

The relationship between field reference plots and variables derived from remote sensing data is of key importance for the performance of the AGB estimation model. Some earlier studies reported an improvement in AGB model performance with increasing plot size, using different

remote sensing data such as LiDAR (Mauya et al., 2015; Zolkos et al., 2013), Radar (Saatchi et al., 2011), or both LiDAR and Radar (Næsset et al., 2015).

Results in this study show that the performance of AGB models derived from the field and the RapidEye data was influenced by the change of field plot size, where the information from remotely sensed data was extracted, and also by the plot shape. The results from the AGB models imply that the model prediction error decreases with the increase in field plot size. However, this result was not observed for Adj R^2.

In this study, an oscillation pattern of the model Adj R^2 was found as plot size changed. This pattern was slightly different than the result from Mauya et al., (2015), which found a general improvement of the model Adj R^2 as the plot size increased from 200 m^2 to 3,000 m^2 and the sample size increased from n=22 to n=30. Different results between the two studies likely arise from the different remote sensing data used. While optical remote sensing data can capture only the variability at canopy surface from above, LiDAR data is able to characterize the vertical stand structures, and forest canopy in 3-dimensional structures. The difference can also be attributed to the range of the tested plot sizes and sample sizes. The range of plot sizes was wider in the current study than the one reported by Mauya et al. (2015).

Increasing field plot sizes tended to increase the model prediction accuracy. This decreasing trend of the model *RMSE* with the increase of plot size has also been reported by earlier studies (Gobakken and Næsset, 2008; Mauya et al., 2015; Næsset et al., 2015; Saatchi et al., 2011; Zolkos et al., 2013). Larger plots are likely to capture more of the variability within the plots and reduce the variance among plots. In other words, the mean estimate values between the plots are more similar to one another and closer to the estimated value of the population mean (Hansen et al., 2015). The declining trend in the model *RMSE* of larger plots is also attributed to the spatial averaging of the error (Mauya et al., 2015; Næsset et al., 2013; Tomppo et al., 2017). Comparisons between the prediction accuracy from Sebangau and Hainich-AWF in the previous sub-study indicated that this accuracy was also attributed to the heterogeneity of the forest stand structures and the AGB variability between sample plots in a dataset.

The square plot shape, such as that used in the EPs, is more influenced by the boundary effect compared to the circular plot shape. Moreover, smaller plots are more influenced by the boundary effects since the ratio between plot perimeter and plot area is larger (Sala and Austin, 2000). In this study, different performances of the AGB models derived from a comparable field plot size of different side lengths were also found. The model derived using rectangular reference plots with 20 m length tended to have lower Adj R^2 compared to the

models derived from comparable-sized rectangular plots with wider side lengths. Furthermore, this study found a significant correlation between the plot perimeter and model prediction accuracy. This correlation was higher than that from plot size. Similarly, a higher correlation was found between the plot perimeter and the model relative maximum residual, compared to that between the plot size and the relative maximum residual. Between the square and the rectangular plot shape, the rectangular plot shape exhibited higher correlation between plot perimeter and Adj R^2 and between plot size and Adj R^2. This result can be explained by the mismatch between the boundary of the plots and the pixels. This effect would be larger in rectangular plots and in smaller plots. However, in the context of forest inventory, a smaller plot is preferable for its efficiency and a rectangular plot shape is commonly used to capture more variability.

In this sub-study, the inclusion of RapidEye data with the Hainich-EP dataset (n=26) improved the precision of the AGB estimation, compared to field observations alone. The highest efficiency from RapidEye data was found to be 3.7 for Dataset 1 and 3.4 for Dataset 2. Both results are considerably high. The study of Næsset et al. (2016) reported an efficiency of 3.3 from RapidEye, which was higher than the efficiency from the InSAR data. They reported that the quality of the image, e.g. free of clouds conditions, was one of the possible reasons for the high efficiency from RapidEye. Indeed, the RapidEye images used over the Hainich study area were also free from clouds.

Based on the results of the current study, the 55x55 m plot size from Dataset 1 or the 50x33.3 m plot size from Dataset 2 were found to be the optimum size for estimating the AGB of the Hainich forest using field and RapidEye data. However, selection of an optimum plot size is always a compromise between the most efficient field cost (time, labor, money) and the target of precision. The use of larger plot size (e.g. plots greater than 1200 m^2 (Mauya et al., 2015) and plots greater than 2500 m^2 (Saatchi et al., 2011)) has been recommended in previous studies that combine field and remote sensing data Regarding the field and the cost efficiencies, an optimum plot size depends also on the forest structure and conditions. For example, based on the experiences in the Sebangau, establishing and measuring large sample plots in the Sebangau forest would be very costly due to difficult access and forest conditions.

In the analysis of this study, it was assumed that the EP dataset followed a SRS sampling design. The availability of a probability sample is the basic requirement to make an effective comparison of the precision inferences (Mauya et al., 2015; Næsset et al., 2015), as it was presented here. Assuming a SRS design will most likely introduce bias in the estimator at an unknown level. However, the EPs were the best data available during the study period for the

assessment of the effect of plot sizes. The sample size of EP was also limited (n=26). Hence, it was also assumed that the results from the limited sample size in this study could provide meaningful information on the effect of different plot sizes in the AGB model derived from the optical remote sensing data.

Chapter 6

Conclusions

One of the requirements for REDD+ country participants is to establish a robust and transparent national forest monitoring system (NFMS). In the implementation of the system, forest inventories and remote sensing are integrated and used as the primary data source. The critical role of remote sensing data for observing, estimating, and monitoring forest biomass is widely recognized, particularly in forests with limited accessibility, as in many tropical countries. The data are also becoming more readily available and accessible to the public, at lower or even no cost. Yet, biomass assessments based on remote sensing data still rely on accurate field reference plots. This study provides a methodological basis of the integration between forest inventories and remote sensing data using the case of the Sebangau tropical peat-swamp forest and the optical remote sensing data of RapidEye.

Regarding the effect of forest restricted visibility in the basal area estimates from ACP, this study demonstrates how to derive the distance of visibility of a forest for implementing ACP method. This information is of importance in choosing a suitable BAF to minimize the effect of forest visibility conditions in ACP estimation. When the distance from an ACP sample point to the target tree exceeds this maximum distance, it leads to an underestimation. The systematic underestimation from ACP with lower BAF is found to be an indication of the effect from forest restricted visibility. A simple approach to test the visibility condition of a forest in the field is by estimating the basal area from a sample point using different BAFs. When the estimation value keeps increasing in large steps with higher BAFs, there is an indication of limited visibility of the sample point. A practical selection of a proper BAF to minimize this effect is the smallest BAF, among two or several BAFs, that gives a similar basal area estimate.

The determination of an optimum BAF in ACP is as necessary as the determination of an optimum plot size in FAP. Selection of a BAF depends on the complexity and the density of the forest stand structures and the density of the ground vegetation. The selected BAF should be the lowest BAF that is least affected by the forest restricted visibility and the highest BAF that provides the highest precision of estimation.

Even though this study was conducted for the basal area estimation of a dense tropical peat-swamp forest, the results also contribute to the implementation of ACP method for estimating other forest variables and for other forest types. The information on the visibility condition of a forest is also useful for the selection between the FAP or ACP plot design.

Regarding research into the integration of remote sensing and field data, this study confirmed the limitation of optical remote sensing to determine the variability of forest AGB, as reported in previous studies, since the data can mainly capture the reflectance from the canopy surfaces. Forests with higher variability of AGB, or forests with higher heterogeneity of stand structures and high diversity of tree species, produce more complex reflectances at the canopy level. Hence, optical remote sensing data is saturated. The red-edge band of RapidEye tends to be superior to the other spectral bands in capturing this variability. In the forest with complex stand structures, GLCM texture features are important predictors for estimating AGB.

Stratification based on the leaf structures, such as coniferous versus broad-leaved, improve the AGB model performance, depending on the heterogeneity of the forest stands structure and plot size. In high heterogeneous stand structures, the variability of the reflectances of the remote sensing data at the canopy level is highly influenced by the shadow effect, thereby weakening the AGB model performance. The AGB estimation model from forests with more homogenous stand structures dominated by trees of relatively small *DBH* may have a higher performance than the model derived from forests with high heterogeneity of stand structures. The sample size of the field data also affects the significant relationships between remote sensing predictor variables and the field AGB but does not necessarily increase the relationship.

Another conclusion from this study is the potential use of RapidEye and field data for estimating forest AGB. The inclusion of RapidEye data offers an opportunity to improve the precision of forest AGB estimates compared to the estimates solely from field observations. The use of model-assisted estimators allows for conducting comparison analyses on the gain in precision from RapidEye in different forest types of different datasets.

The plot size and plot shape have a great effect on the performance of AGB model derived from RapidEye and field data and on the model-assisted AGB estimation. Analysis of the effect of plot size and plot shape is of importance for survey design, and plot size and shape are among the most important design elements of field surveys.

This study confirmed the findings of earlier studies regarding the contribution of remote sensing data to improving the precision of carbon estimation. However, to obtain a higher

efficiency from combining field and remote sensing data to improve the precision of estimation, selection of plot size and shape should also consider remotely sensed data. An optimum size of field inventory plots might not be large enough to enable the remotely sensed data to capture all variability within the plots. A perfect co-registration between the plot and the pixel boundary is hard to achieve. A plot shape of larger perimeter with the same area is likely to increase the mismatch between the two boundaries and reduce the model performance. Thus, it is not recommended.

Choosing an optimum plot size should also consider the forest stand structures. In more homogenous forest stands, the optimum plot size could be smaller than in a forest with high heterogeneity of the stand structures.

Results from the current study contribute to the Indonesian government's national program for estimating the amount of carbon from forests to assess the country's commitment to reduce carbon emissions from the forestry sector. Many of the Indonesian forests are still difficult to access and are poor in infrastructure. Since the available budget is always limited, an effective survey design is required, in which the potential of remote sensing data can be leveraged to enhance the estimation.

The results are also relevant for the plan of the Indonesian government to re-designing and to improve the NFI to include optical remote sensing data of higher spatial resolution than Landsat for the calculation of the national FREL. In this context, RapidEye can be considered as an option to be implemented in the carbon estimation of the Indonesian forests. The RapidEye constellation of 5 satellites increases the possibility of having more acquisition data, allowing for the formation of a mosaic image with minimum cloud coverage.

References

Albert, K., Annighöfer, P., Schumacher, J., Ammer, C., 2014. Biomass equations for seven different tree species growing in coppice-with-standards forests in Central Germany. Scandinavian Journal of Forest Research 29, 210–221. https://doi.org/10.1080/02827581.2014.910267

Anaya, J.A., Chuvieco, E., Palacios-Orueta, A., 2009. Aboveground biomass assessment in Colombia: A remote sensing approach. Forest Ecology and Management 257, 1237–1246. https://doi.org/10.1016/j.foreco.2008.11.016

Armon, R.H., Hänninen, O., 2015. Environmental Indicators. Springer.

Austin, K.G., Mosnier, A., Pirker, J., McCallum, I., Fritz, S., Kasibhatla, P.S., 2017. Shifting patterns of oil palm driven deforestation in Indonesia and implications for zero-deforestation commitments. Land Use Policy 69, 41–48. https://doi.org/10.1016/j.landusepol.2017.08.036

Avery, T.E., Burkhart, H.E., 2002. Forest measurements, 5th Edition. ed. McGraw-Hill.

Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P.S.A., Dubayah, R., Friedl, M.A., Samanta, S., Houghton, R.A., 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature Climate Change 2, 182–185. https://doi.org/10.1038/nclimate1354

Baccini, A., Laporte, N., Goetz, S.J., Sun, M., Dong, H., 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. Environ. Res. Lett. 3, 045011. https://doi.org/10.1088/1748-9326/3/4/045011

Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D., Houghton, R.A., 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. Science 358, 230–234. https://doi.org/10.1126/science.aam5962

Ballhorn, U., Jubanski, J., Siegert, F., 2011. ICESat/GLAS Data as a Measurement Tool for Peatland Topography and Peat Swamp Forest Biomass in Kalimantan, Indonesia. Remote Sensing 3, 1957–1982. https://doi.org/10.3390/rs3091957

Bannari, A., Morin, D., Bonn, F., Huete, A.R., 1995. A review of vegetation indices. Remote Sensing Reviews 13, 95–120. https://doi.org/10.1080/02757259509532298

Barnes, E.M., Clarke, T.R., Richards, S.E., Colaizzi, P.D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., others, 2000. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data, in: Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, MN. pp. 16–19.

Becker, P., Nichols, T., 2011. Effects of Basal Area Factor and Plot Size on Precision and Accuracy of Forest Inventory Estimates. Northern Journal of Applied Forestry 28, 152–156.

Birkhofer, K., Schöning, I., Alt, F., Herold, N., Klarner, B., Maraun, M., Marhan, S., Oelmann, Y., Wubet, T., Yurkov, A., Begerow, D., Berner, D., Buscot, F., Daniel, R., Diekötter, T., Ehnes, R.B., Erdmann, G., Fischer, C., Foesel, B., Groh, J., Gutknecht, J., Kandeler, E., Lang, C., Lohaus, G., Meyer, A., Nacke, H., Näther, A., Overmann, J., Polle, A., Pollierer, M.M., Scheu, S., Schloter, M., Schulze, E.-D., Schulze, W., Weinert, J., Weisser, W.W., Wolters, V., Schrumpf, M., 2012. General Relationships between Abiotic Soil Properties and Soil Biota across Spatial Scales and Different Land-Use Types. PLOS ONE 7, e43292. https://doi.org/10.1371/journal.pone.0043292

Bitterlich, W., 1984. The relascope idea. Relative measurements in forestry. Commonwealth Agricultural Bureaux, Farnham Royal.

BlackBidge, 2015. RapidEye: Satellite Imagery Product Specifications. Version 6.1. [WWW Document]. URL http://www.e-geos.it/images/Satellite_data/RAPIDEYE/RE_Product_Specifications_ENG.pdf

BlackBridge, n.d. White Paper: The RapidEye Red Edge Band [WWW Document]. URL https://resa.blackbridge.com/files/2014-06/Red_Edge_White_Paper.pdf (accessed 4.13.15).

BMKG, 2014. Prakiraan Musim Hujan 2014/2015 di Indonesia. Badan Meteorologi, Klimatologi, dan Geofisika. Jakarta, 2014.

Bojanowski, J., 2007. The analysis of sensibility to the change of the input parameters in the 6S model. Millpress, Rotterdam.

Bonan, G.B., 2008. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. Science 320, 1444–1449. https://doi.org/10.1126/science.1155121

Box, G.E.P., Cox, D.R., 1964. An Analysis of Transformations. Journal of the Royal Statistical Society. Series B (Methodological) 26, 211–252.

Boyd, D.S., Danson, F.M., 2005. Satellite remote sensing of forest resources: three decades of research development. Progress in Physical Geography: Earth and Environment 29, 1–26. https://doi.org/10.1191/0309133305pp432ra

Brooks, J.R., 2006. An evaluation of big basal area factor sampling in Appalachian Hardwoods. Northern Journal of Applied Forestry 23, 62–65.

Brooks, J.R., McGill, D.W., 2004. Evaluation of multiple fixed area plot sizes and BAFs in even-aged hardwood stands.

Brown, S., 1997. Estimating Biomass and Biomass Change of Tropical Forests: A Primer. FAO, Rome, Italy.

Campbell, L.A.D., 2013. Bornean toprical peat swamp forests Disturbance effects on carbon content and tree species traits in tropical peat swamp forest (Bsc. Thesis). Dalhousie University.

Canty, A., Ripley, B., 2016. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-18, https://CRAN.R-project.org/package=boot.

Castillo-Santiago, M.A., Ricker, M., de Jong, B.H.J., 2010. Estimation of tropical forest structure from SPOT-5 satellite images. International Journal of Remote Sensing 31, 2767–2782. https://doi.org/10.1080/01431160903095460

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B.W., Ogawa, H., Puig, H., Riéra, B., Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145, 87–99. https://doi.org/10.1007/s00442-005-0100-x

Cronin, T., Santoso, L., 2010. REDD+ politics in the media: a case study from Indonesia. Center for International Forestry Research. https://doi.org/10.17528/cifor/003275

Eckert, S., 2012. Improved Forest Biomass and Carbon Estimations Using Texture Measures from WorldView-2 Satellite Data. Remote Sensing 4, 810–829. https://doi.org/10.3390/rs4040810

Eichkitz, C.G., Davies, J., Amtmann, J., Schreilechner, M.G., de Groot, P., 2015. Grey level co-occurrence matrix and its application to seismic data. first break 33, 71–77.

Eitel, J.U., Vierling, L.A., Litvak, M.E., Long, D.S., Schulthess, U., Ager, A.A., Krofcheck, D.J., Stoscheck, L., 2011. Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. Remote Sensing of Environment 115, 3640–3646. https://doi.org/10.1016/j.rse.2011.09.002

Englhart, S., Franke, J., Keuck, V., Siegert, F., 2012a. Aboveground biomass estimation of tropical peat swamp forests using SAR and optical data, in: 2012 IEEE International Geoscience and Remote Sensing Symposium. Presented at the 2012 IEEE International Geoscience and Remote Sensing Symposium, pp. 6577–6580. https://doi.org/10.1109/IGARSS.2012.6352092

Englhart, S., Jubanski, J., Siegert, F., 2013. Quantifying Dynamics in Tropical Peat Swamp Forest Biomass with Multi-Temporal LiDAR Datasets. Remote Sensing 5, 2368–2388. https://doi.org/10.3390/rs5052368

Englhart, S., Keuck, V., Siegert, F., 2012b. Modeling Aboveground Biomass in Tropical Forests Using Multi-Frequency SAR Data - A Comparison of Methods. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5, 298–306. https://doi.org/10.1109/JSTARS.2011.2176720.

Englhart, S., Keuck, V., Siegert, F., 2011. Aboveground biomass retrieval in tropical forests — The potential of combined X- and L-band SAR data use. Remote Sensing of Environment 115, 1260–1271. https://doi.org/10.1016/j.rse.2011.01.008.

Erkkilä, A., Siiskonen, H., 1992. Forestry in Namibia 1850-1990, Silva Carelica. University of Joensuu, Finland.

FAO, 2016. Global forest resources assessment 2015: how are the world's forests changing?, Second. ed. FAO, Rome, Italy.

FAO, 2015. Global forest resources assessment 2015. Desk reference. FAO, Rome, Italy.

FAO, 2010. Global Forest Resources Assessment 2010. Main Report. FAO, Rome, Italy.

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh, A., 2016. Review of studies on tree species classification from remotely sensed data. Remote Sensing of Environment 186, 64–87. https://doi.org/10.1016/j.rse.2016.08.013

Fehrmann, L., 2007. Alternative Methoden zur Biomasseschätzung auf Einzelbaumebene unter spezieller Berücksichtigung der k-Nearest Neighbour (k-NN) Methode. Göttingen.

Fischer, C., Kleinn, C., Fehrmann, L., Fuchs, H., Panferov, O., 2011. A national level forest resource assessment for Burkina Faso – A field based forest inventory in a semiarid environment combining small sample size with large observation plots. Forest Ecology and Management 262, 1532–1540. https://doi.org/10.1016/j.foreco.2011.07.001

Fischer, M., Bossdorf, O., Gockel, S., Hänsel, F., Hemp, A., Hessenmöller, D., Korte, G., Nieschulze, J., Pfeiffer, S., Prati, D., Renner, S., Schöning, I., Schumacher, U., Wells, K., Buscot, F., Kalko, E.K.V., Linsenmair, K.E., Schulze, E.-D., Weisser, W.W., 2010. Implementing large-scale and long-term functional biodiversity research: The Biodiversity Exploratories. Basic and Applied Ecology 11, 473–485. https://doi.org/10.1016/j.baae.2010.07.009

Fox, J., Weisberg, S., 2011. An R Companion to Applied Regression, Second. ed. Sage, Thousand Oaks CA. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Fuchs, H., Magdon, P., Kleinn, C., Flessa, H., 2009. Estimating aboveground carbon in a catchment of the Siberian forest tundra: Combining satellite imagery and field inventory. Remote Sensing of Environment 113, 518–531. https://doi.org/10.1016/j.rse.2008.07.017

Gallaun, H., Zanchi, G., Nabuurs, G.-J., Hengeveld, G., Schardt, M., Verkerk, P.J., 2010. EU-wide maps of growing stock and above-ground biomass in forests based on remote sensing and field measurements. Forest Ecology and Management 260, 252–261. https://doi.org/10.1016/j.foreco.2009.10.011

Gitelson, A., Merzlyak, M.N., 1994. Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. Journal of Plant Physiology 143, 286–292. https://doi.org/10.1016/S0176-1617(11)81633-0

Gitelson, A.A., Merzlyak, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. International Journal of Remote Sensing 18, 2691–2697. https://doi.org/10.1080/014311697217558

Gobakken, T., Næsset, E., 2009. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. Can. J. For. Res. 39, 1036–1052. https://doi.org/10.1139/X09-025

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, 1095–1109. https://doi.org/10.1139/X07-219

Goetz, S.J., Baccini, A., Laporte, N.T., Johns, T., Walker, W., Kellndorfer, J., Houghton, R.A., Sun, M., 2009. Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. Carbon Balance and Management 4, 2. https://doi.org/10.1186/1750-0680-4-2

Grosenbaugh, L.R., 1958. Point-Sampling and Line-Sampling Probability Theory, Geometric Implications, Synthesis, Southern Forest Experiment Station, Occasional Paper 160. Forest Service, U.S. Dept. of Agriculture, New Orleans.

Hall-Beyer, M., 2007. GLCM Texture Tutorial [WWW Document]. URL http://www.fp.ucalgary.ca/mhallbey/tutorial.htm (accessed 4.3.17).

Hansen, E.H., Gobakken, T., Bollandsås, O.M., Zahabu, E., Næsset, E., 2015. Modeling Aboveground Biomass in Dense Tropical Submontane Rainforest Using Airborne Laser Scanner Data. Remote Sensing 7, 788–807. https://doi.org/10.3390/rs70100788

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics SMC-3, 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Harrell Jr, F.E., Dupont, with contributions from C., others, many, 2016. Hmisc: Harrell Miscellaneous. https://CRAN.R-project.org/package=Hmisc.

Hasenauer, H., Eastaugh, C.S., 2012. Assessing forest production using terrestrial monitoring data. International Journal of Forestry Research 2012. http://dx.doi.org/10.1155/2012/961576

Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling. https://CRAN.R-project.org/package=raster.

Hirano, T., Kusin, K., Limin, S., Osaki, M., 2014. Carbon dioxide emissions through oxidative peat decomposition on a burnt tropical peatland. Global Change Biology 20, 555–565. https://doi.org/10.1111/gcb.12296

Hirano, T., Segah, H., Kusin, K., Limin, S., Takahashi, H., Osaki, M., 2012. Effects of disturbances on the carbon balance of tropical peat swamp forests. Global Change Biology 18, 3410–3422. https://doi.org/10.1111/j.1365-2486.2012.02793.x

Hooijer, A., Page, S., Canadell, J.G., Silvius, M., Kwadijk, J., Wösten, H., Jauhiainen, J., 2010. Current and future CO2 emissions from drained peatlands in Southeast Asia. Biogeosciences 7, 1505–1514. https://doi.org/10.5194/bg-7-1505-2010

Hooijer, A., Silvius, M., Wösten, H., Page, S., 2006. PEAT-CO2: assessment of CO2 emissions from drained peatlands in SE Asia, 1st edition. Delft Hydraulics report Q3943 (2006).

Houghton, R.A., 1991. Releases of carbon to the atmosphere from degradation of forests in tropical Asia. Can. J. For. Res. 21, 132–142. https://doi.org/10.1139/x91-017

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote sensing of environment 83, 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

Husch, B., Beers, T.W., Kershaw, Jr, J.A., 1982. Forest mensuration. John Wiley & Sons.

Indonesia MoF, 2004. Keputusan Menhut No. SK.423/Menhut-II/2004.

Indonesia MoF, 1992. Langkah-langkah prosedur sampling lapangan untuk proyek inventarisasi hutan nasional.

IPCC, 2014. Climate change 2014: mitigation of climate change: Working Group III contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, New York. http://www.ipcc.ch/report/ar5/wg3/.

Jackson, R.D., Huete, A.R., 1991. Interpreting vegetation indices. Preventive Veterinary Medicine 11, 185–200. https://doi.org/10.1016/S0167-5877(05)80004-2

Jaenicke, J., Rieley, J.O., Mott, C., Kimman, P., Siegert, F., 2008. Determination of the amount of carbon stored in Indonesian peatlands. Geoderma 147, 151–158. https://doi.org/10.1016/j.geoderma.2008.08.008

Jochem, A., Hollaus, M., Rutzinger, M., Höfle, B., 2011. Estimation of aboveground biomass in alpine forests: a semi-empirical approach considering canopy transparency derived from airborne LiDAR data. Sensors (Basel) 11, 278–295. https://doi.org/10.3390/s110100278

Jones, H.G., Vaughan, R.A., 2010. Remote Sensing of Vegetation: Principles, Techniques, and Applications. OUP Oxford.

Joshi, N., Mitchard, E.T., Woo, N., Torres, J., Moll-Rocek, J., Ehammer, A., Murray Collins, Jepsen, M.R., Fensholt, R., 2015. Mapping dynamics of deforestation and forest degradation in tropical forests using radar satellite data. Environ. Res. Lett. 10, 034014. https://doi.org/10.1088/1748-9326/10/3/034014

Jubanski, J., Ballhorn, U., Kronseder, K., J Franke, Siegert, F., 2013. Detection of large above-ground biomass variability in lowland forest ecosystems by airborne LiDAR. Biogeosciences 10, 3917–3930. https://doi.org/10.5194/bg-10-3917-2013

Kharazipour, A.R., Schöpper, C., Müller, C., Euring, M., 2009. Review of forests, wood products and wood biotechnology of Iran and Germany - Part III. https://doi.org/10.17875/gup2009-272

Köhl, M., Magnussen, S., Marchetti, M., 2006. Sampling methods, remote sensing and GIS multiresource forest inventory, Tropical forestry. Springer, Berlin ; London.

Kramer, H., Akça, A., 1995. Leitfaden zur Waldmeßlehre. Sauerländer Verlag, Frankfurt am Main.

Krieger, D.J., 2001. Economic Value of Forest Ecosystem Services: A Review. The Wilderness Society, Washington, D.C. https://www.sierraforestlegacy.org/Resources/Conservation/FireForestEcology/ForestEconomics/EcosystemServices.pdf.

Kronseder, K., Ballhorn, U., Böhm, V., Siegert, F., 2012. Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan using LiDAR data. International Journal of Applied Earth Observation and Geoinformation 18, 37–48. https://doi.org/10.1016/j.jag.2012.01.010

Latifi, H., Nothdurft, A., Straub, C., Koch, B., 2012. Modelling stratified forest attributes using optical/LiDAR features in a central European landscape. International Journal of Digital Earth 5, 106–132. https://doi.org/10.1080/17538947.2011.583992

Lu, D., 2005. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. International Journal of Remote Sensing 26, 2509–2525. https://doi.org/10.1080/01431160500142145

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 9, 63–105. https://doi.org/10.1080/17538947.2014.990526

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., Vaglio Laurin, G., Saah, D., 2012. Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. International Journal of Forestry Research 2012, 16. http://dx.doi.org/10.1155/2012/436537

Lu, D., Mausel, P., Brondízio, E., Moran, E., 2004. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. Forest Ecology and Management 198, 149–167. https://doi.org/10.1016/j.foreco.2004.03.048

Magdon, P., González-Ferreiro, E., Pérez-Cruzado, C., Purnama, E.S., Sarodja, D., Kleinn, C., 2018. Evaluating the Potential of ALS Data to Increase the Efficiency of Aboveground Biomass Estimates in Tropical Peat–Swamp Forests. Remote Sensing 10, 1344. https://doi.org/10.3390/rs10091344

Manuri, S., Brack, C., Nugroho, N.P., Hergoualc'h, K., Novita, N., Dotzauer, H., Verchot, L., Putra, C.A.S., Widyasari, E., 2014. Tree biomass equations for tropical peat swamp forest ecosystems in Indonesia. Forest Ecology and Management 334, 241–253. https://doi.org/10.1016/j.foreco.2014.08.031

Mascaro, J., Detto, M., Asner, G.P., Muller-Landau, H.C., 2011. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. Remote Sensing of Environment 115, 3770–3774. https://doi.org/10.1016/j.rse.2011.07.019

Masripatin, N. (Ed.), 2017. Third National Communication Indonesia under the UNFCCC. Ministry of Environment and Forestry, Indonesia.

Mauya, E.W., Hansen, E.H., Gobakken, T., Bollandsås, O.M., Malimbwi, R.E., 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, 10. https://doi.org/10.1186/s13021-015-0021-x

McRoberts, R.E., Chen, Q., Domke, G.M., Ståhl, G., Saarela, S., Westfall, J.A., 2016. Hybrid estimators for mean aboveground carbon per unit area. Forest Ecology and Management 378, 44–56. https://doi.org/10.1016/j.foreco.2016.07.007

McRoberts, R.E., Næsset, E., Gobakken, T., 2013a. Inference for lidar-assisted estimation of forest growing stock volume. Remote Sensing of Environment 128, 268–275. https://doi.org/10.1016/j.rse.2012.10.007

McRoberts, R.E., Tomppo, E.O., Vibrans, A.C., de Freitas, J.V., 2013b. Design Considerations for Tropical Forest Inventories. https://doi.org/10.4336/2013.pfb.33.74.430

Meister, K., Ashton, M.S., Craven, D., Griscom, H., 2012. Carbon Dynamics of Tropical Forests, in: Managing Forest Carbon in a Changing Climate. Springer, Dordrecht, pp. 51–75. https://doi.org/10.1007/978-94-007-2232-3_4

Mendelsohn, J., Jarvis, A., Roberts, C., Robertson, T., 2002. Atlas of Namibia: A portrait of the land and its people. David Philip, Cape Town. https://doi.org/10.1002/mmnz.20040800111

Miller, T.L. using F. code by A., 2009. leaps: regression subset selection. https://CRAN.R-project.org/package=leaps.

MoEF, 2016. National Forest Reference Emission Level for Deforestation and Forest Degradation: In the Context of Decision 1/CP.16 para 70 UNFCCC (Encourages developing country Parties to contribute to mitigation actions in the forest sector), Directorate General of Climate Change. The Ministry of Environment and Forestry. Indonesia.

Mölder, A., Bernhardt-Römermann, M., Schmidt, W., 2006. Forest ecosystem research in Hainich National Park (Thuringia): First results on flora and vegetation in stands with contrasting tree species diversity. Waldökologie Online 3, 83–99.

Molto, Q., Rossi, V., Blanc, L., 2013. Error propagation in biomass estimation in tropical forests. Methods in Ecology and Evolution 4, 175–183. https://doi.org/10.1111/j.2041-210x.2012.00266.x

Morrogh-Bernard, H.C.C., 2009. Orang-utan Behavioural Ecology in the Sabangau Pear-swamp Forest, Borneo (PhD. Thesis). University of Cambridge.

Motz, K., Sterba, H., Pommerening, A., 2010. Sampling measures of tree diversity. Forest Ecology and Management 260, 1985–1996. https://doi.org/10.1016/j.foreco.2010.08.046

Mutanga, O., Adam, E., Cho, M.A., 2012. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. International Journal of Applied Earth Observation and Geoinformation 18, 399–406. https://doi.org/10.1016/j.jag.2012.03.012

Myneni, R.B., Hall, F.G., Sellers, P.J., Marshak, A.L., 1995. The interpretation of spectral vegetation indexes. IEEE Transactions on Geoscience and Remote Sensing 33, 481–486. https://doi.org/10.1109/36.377948

Næsset, E., Bollandsås, O.M., Gobakken, T., Gregoire, T.G., Ståhl, G., 2013. Model-assisted estimation of change in forest biomass over an 11year period in a sample survey supported by airborne LiDAR: A case study with post-stratification to provide "activity data." Remote Sensing of Environment 128, 299–314. https://doi.org/10.1016/j.rse.2012.10.008

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. https://doi.org/10.1016/j.rse.2015.07.002

Næsset, E., Ørka, H.O., Solberg, S., Bollandsås, O.M., Hansen, E.H., Mauya, E., Zahabu, E., Malimbwi, R., Chamuya, N., Olsson, H., Gobakken, T., 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: A comparison of estimated precision. Remote Sensing of Environment 175, 282–300. https://doi.org/10.1016/j.rse.2016.01.006

Namibia MoAWF, 2006. Forest act, 2001: Declaration of an area as a community forest: Ncaute Community Forest.

O'brien, R.M., 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity 41, 673–690. https://doi.org/10.1007/s11135-006-9018-6

Ollinger, S.V., 2011. Sources of variability in canopy reflectance and the convergent properties of plants. New Phytologist 189, 375–394. https://doi.org/10.1111/j.1469-8137.2010.03536.x

Osaki, M., Setiadi, B., Takahashi, H., Evri, M., 2016. Peatland in Kalimantan, in: Osaki, M., Tsuji, N. (Eds.), Tropical Peatland Ecosystems. Springer Japan, Tokyo, pp. 91–112. https://doi.org/10.1007/978-4-431-55681-7_6

Page, S.E., Hooijer, A., 2016. In the line of fire: the peatlands of Southeast Asia. Phil. Trans. R. Soc. B 371, 20150176. https://doi.org/10.1098/rstb.2015.0176

Page, S.E., Rieley, J.O., Banks, C.J., 2011. Global and regional importance of the tropical peatland carbon pool. Global Change Biology 17, 798–818. https://doi.org/10.1111/j.1365-2486.2010.02279.x

Page, S.E., Rieley, J.O., Shotyk, Ø.W., Weiss, D., 1999. Interdependence of peat and vegetation in a tropical peat swamp forest. Philosophical Transactions of the Royal Society of London B: Biological Sciences 354, 1885–1897. https://doi.org/10.1098/rstb.1999.0529

Page, S.E., Siegert, F., Rieley, J.O., Boehm, H.-D.V., Jaya, A., Limin, S., 2002. The amount of carbon released from peat and forest fires in Indonesia during 1997. Nature 420, 61–65. https://doi.org/10.1038/nature01131

PEACE, 2007. Indonesia and Climate Change: Current Status and Policies.

Peel, M.C., Finlayson, B.L., Mcmahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrology and Earth System Sciences Discussions 4, 439–473.

Pérez-Cruzado, C., Fehrmann, L., Magdon, P., Cañellas, I., Sixto, H., Kleinn, C., 2015. On the site-level suitability of biomass models. Environmental Modelling & Software 73, 14–26. https://doi.org/10.1016/j.envsoft.2015.07.019

Pérez-Cruzado, C., Merino, A., Rodríguez-Soalleiro, R., 2011. A management tool for estimating bioenergy production and carbon sequestration in Eucalyptus globulus and Eucalyptus nitens grown as short rotation woody crops in north-west Spain. biomass and bioenergy 35, 2839–2851. https://doi.org/10.1016/j.biombioe.2011.03.020

Pfeifer, M., Kor, L., Nilus, R., Turner, E., Cusack, J., Lysenko, I., Khoo, M., Chey, V.K., Chung, A.C., Ewers, R.M., 2016. Mapping the structure of Borneo's tropical forests across a degradation gradient. Remote Sensing of Environment 176, 84–97. https://doi.org/10.1016/j.rse.2016.01.014

Picard, N., Bosela, F.B., Rossi, V., 2015. Reducing the error in biomass estimates strongly depends on model selection. Annals of Forest Science 72, 811–823. https://doi.org/10.1007/s13595-014-0434-9

Picard, N., Henry, M., Fonton, N.H., Kondaoulé, J., Fayolle, A., Birigazzi, L., Sola, G., Poultouchidou, A., Trotta, C., Maïdou, H., 2016. Error in the estimation of emission factors for forest degradation in central Africa. Journal of Forest Research 21, 23–30. https://doi.org/10.1007/s10310-015-0510-5

Piqué, M., Obon, B., Condés, S., Saura, S., 2011. Comparison of relascope and fixed-radius plots for the estimation of forest stand variables in northeast Spain: an

inventory simulation approach. European Journal of Forest Research 130, 851–859. https://doi.org/10.1007/s10342-010-0477-x

Proisy, C., Couteron, P., Fromard, F., 2007. Predicting and mapping mangrove biomass from canopy grain analysis using Fourier-based textural ordination of IKONOS images. Remote Sensing of Environment 109, 379–392. https://doi.org/10.1016/j.rse.2007.01.009

Putra, E.I., Hayasaka, H., Takahashi, H., Usup, A., 2008. Recent Peat Fire Activity in the Mega Rice Project Area, Central Kalimantan, Indonesia. Journal of Disaster Research Vol.3 No.5 9.

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sensing of Environment 48, 119–126. https://doi.org/10.1016/0034-4257(94)90134-1

R Core Team, 2015. R: A Language and Environment for Statistical Computing [WWW Document]. URL http://www.R-poject.org/

Rana, P., Tokola, T., Korhonen, L., Xu, Q., Kumpula, T., Vihervaara, P., Mononen, L., 2013. Training Area Concept in a Two-Phase Biomass Inventory Using Airborne Laser Scanning and RapidEye Satellite Data. Remote Sensing 6, 285–309. https://doi.org/10.3390/rs6010285

Ritter, T., Nothdurft, A., Saborowski, J., 2013. Correcting the nondetection bias of angle count sampling. Can. J. For. Res. 43, 344–354. https://doi.org/10.1139/cjfr-2012-0408

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. Third ERTS-1 Symposium NASA, NASA SP-351, Washington DC. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022614.pdf.

Rutishauser, E., Noor'an, F., Laumonier, Y., Halperin, J., Rufi'ie, Hergoualc'h, K., Verchot, L., 2013. Generic allometric models including height best estimate forest biomass and carbon stocks in Indonesia. Forest Ecology and Management 307, 219–225. https://doi.org/10.1016/j.foreco.2013.07.013

Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., Hyyppä, J., 2015. Model-assisted estimation of growing stock volume using different combinations of LiDAR and Landsat data as auxiliary information. Remote Sensing of Environment 158, 431–440. https://doi.org/10.1016/j.rse.2014.11.020

Saarela, S., Schnell, S., Tuominen, S., Balázs, A., Hyyppä, J., Grafström, A., Ståhl, G., 2016. Effects of positional errors in model-assisted and model-based estimation of growing stock volume. Remote Sensing of Environment 172, 101–108. https://doi.org/10.1016/j.rse.2015.11.002

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, DESDynI VEG-3D Special Issue 115, 2836–2849. https://doi.org/10.1016/j.rse.2010.07.015

Sala, O.E., Austin, A.T., 2000. Methods of Estimating Aboveground Net Primary Productivity, in: Methods in Ecosystem Science. Springer, New York, NY, pp. 31–43. https://doi.org/10.1007/978-1-4612-1224-9_3

Sarker, L.R., Nichol, J.E., 2011. Improved forest biomass estimates using ALOS AVNIR-2 texture indices. Remote Sensing of Environment 115, 968–977. https://doi.org/10.1016/j.rse.2010.11.010

Schlund, M., von Poncet, F., Kuntz, S., Schmullius, C., Hoekman, D.H., 2015. TanDEM-X data for aboveground biomass retrieval in a tropical peat swamp forest. Remote Sensing of Environment 158, 255–266. https://doi.org/10.1016/j.rse.2014.11.016

Schmidt, A. von, 1967. Der rechnerische Ausgleich von Bestandeshöhenkurven. Forstwissenschaftliches Centralblatt 86, 370–382.

Schreuder, H.T., Rennie, J.C., Williams, M., 1992. Comparison of three sampling schemes for estimating frequency and D2H by diameter class - a simulation study. Forest Ecology and Management 50, 117–131. https://doi.org/10.1016/0378-1127(92)90318-4

Scott, C.T., 1990. An overview of fixed versus variable-radius plots for successive inventories. USDA Forest Service general technical report PNW-GTR - Pacific Northwest Research Station (USA).

Scrinzi, G., Clementel, F., Floris, A., 2015. Angle count sampling reliability as ground truth for area-based LiDAR applications in forest inventories. Can. J. For. Res. 45, 506–514. https://doi.org/10.1139/cjfr-2014-0408

Sibanda, M., Mutanga, O., Rouget, M., 2015. Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. ISPRS Journal of Photogrammetry and Remote Sensing 110, 55–65. https://doi.org/10.1016/j.isprsjprs.2015.10.005

Sileshi, G.W., 2014. A critical review of forest biomass estimation models, common mistakes and corrective measures. Forest Ecology and Management 329, 237–254. https://doi.org/10.1016/j.foreco.2014.06.026

Skole, D.L., Samek, J.H., Smalligan, M.J., 2011. Implications of allometry. Proceedings of the National Academy of Sciences 108, E12. https://doi.org/10.1073/pnas.1015854108

Sorensen, K.W., 1993. Indonesian peat swamp forests and their role as a carbon sink. Chemosphere 27, 1065–1082. https://doi.org/10.1016/0045-6535(93)90068-G

Sterba, H., 2008. Diversity indices based on angle count sampling and their interrelationships when used in forest inventories. Forestry 81, 587–597. https://doi.org/10.1093/forestry/cpn010

Strack, M. (Ed.), 2008. Peatlands and Climate Change. International Peat Society, Jyväskylä, Finland. http://www.peatsociety.org/sites/default/files/files/PeatlandsandClimateChangeBookIPS2008.pdf.

Suganuma, H., Abe, Y., Taniguchi, M., Tanouchi, H., Utsugi, H., Kojima, T., Yamada, K., 2006. Stand biomass estimation method by canopy coverage for application to remote sensing in an arid area of Western Australia. Forest Ecology and Management 222, 75–87. https://doi.org/10.1016/j.foreco.2005.10.014

Sukwong, S., Frayer, W.E., Mogren, E.W., 1971. Generalized Comparisons of the Precision Fixed-Radius and Variable-Radius Plots for Basal-Area Estimates. Forest Science 17, 263–271.

Thenkabail, P.S., Stucky, N., Griscom, B.W., Ashton, M.S., Diels, J., Meer, B. van der, Enclona, E., 2004. Biomass estimations and carbon stock calculations in the oil palm plantations of African derived savannas using IKONOS data. International Journal of Remote Sensing 25, 5447–5472. https://doi.org/10.1080/01431160412331291279

Tomppo, E.O., Kuusinen, N., Mäkisara, K., Katila, M., McRoberts, R.E., 2017. Effects of field plot configurations on the uncertainties of ALS-assisted forest resource estimates. Scandinavian Journal of Forest Research 32, 488–500. https://doi.org/10.1080/02827581.2016.1259425

Tomppo, E.O., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., Katila, M., 2008. Combining national forest inventory field plots and remote sensing data for forest databases. Remote Sensing of Environment, Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue 112, 1982–1999. https://doi.org/10.1016/j.rse.2007.03.032

Torres, A.B., Lovett, J.C., 2012. Using basal area to estimate aboveground carbon stocks in forests: La Primavera Biosphere's Reserve, Mexico. Forestry cps084. https://doi.org/10.1093/forestry/cps084

UNFCCC, 2009. Methodological guidance for activities relating to reducing emissions from deforestation and forest degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries. Decision 4/CP. 15. See h ttp. unfccc. int/resource/docs/2009/cop15/eng/11a01. pdf# page 11.

Urbazaev, M., Thiel, C., Cremer, F., Dubayah, R., Migliavacca, M., Reichstein, M., Schmullius, C., 2018. Estimation of forest aboveground biomass and uncertainties by integration of field measurements, airborne LiDAR, and SAR and optical satellite data in Mexico. Carbon Balance Manage 13, 5. https://doi.org/10.1186/s13021-018-0093-5

Van Laar, A., Akça, A., 2007. Forest Mensuration. Springer Netherlands.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, Fourth. ed. Springer, New York. http://www.stats.ox.ac.uk/pub/MASS4.

Verchot, L.V., Petkova, E., Obidzinski, K., Atmadja, S., Yuliani, L., Dermawan, A., Murdiyarso, D., Amira, S., 2010. Reducing forestry emissions in Indonesia. Center for International Forestry Research, Bogor, Indonesia. https://www.cifor.org/library/3142/reducing-forestry-emissions-in-indonesia/

Vermote, E., Tanré, D., Deuzé, J.L., Herman, M., Morcrette, J.J., Kotchenova, S.Y., 2006. Second simulation of a satellite signal in the solar spectrum-vector (6SV). 6S User Guide Version 3, 1–55.

Wallner, A., Elatawneh, A., Schneider, T., Knoke, T., 2015. Estimation of forest structural information using RapidEye satellite data. Forestry (Lond) 88, 96–107. https://doi.org/10.1093/forestry/cpu032

West, P.W., 2009. Tree and Forest Measurement, Second Edition, 2nd ed. 2009. ed. Springer Berlin Heidelberg, Berlin ; New York.

Wijaya, A., Kusnadi, S., Gloaguen, R., Heilmeier, H., 2010a. Improved strategy for estimating stem volume and forest biomass using moderate resolution remote

sensing data and GIS. Journal of Forestry Research 21, 1–12. https://doi.org/10.1007/s11676-010-0001-7

Wijaya, A., Liesenberg, V., Gloaguen, R., 2010b. Retrieval of forest attributes in complex successional forests of Central Indonesia: Modeling and estimation of bitemporal data. Forest Ecology and Management 259, 2315–2326. https://doi.org/10.1016/j.foreco.2010.03.004

Wulder, M., 1998. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. Progress in Physical Geography 22, 449–476. https://doi.org/10.1177/030913339802200402

Yoo, S., Im., J., Wagner, J.E., 2012. Variable selection for hedonic model using machine learning approaches: A case study in Onondaga County, NY. Landscape and Urban Planning 107, 293–306. https://doi.org/10.1016/j.landurbplan.2012.06.009

Zeide, B., 1985. Distance to point-sampled trees and their diameters. Forest ecology and management 11, 131–137. http://www.sciencedirect.com/science/article/pii/0378112785900623

Zhang, X., Zhao, Y., Ashton, M.S., Lee, X., 2012. Measuring Carbon in Forests, in: Managing Forest Carbon in a Changing Climate. Springer, Dordrecht, pp. 139–164. https://doi.org/10.1007/978-94-007-2232-3_7

Zolkos, S.G., Goetz, S.J., Dubayah, R., 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. Remote Sensing of Environment 128, 289–298. https://doi.org/10.1016/j.rse.2012.10.017

Zvoleff, A., 2016. glcm: Calculate Textures from Grey-Level Co-Occurrence Matrices (GLCMs). https://CRAN.R-project.org/package=glcm.

Appendices

APPENDIX A. Determination of the plot design of LSI and HIL plots

The approach was based on a consideration of having a comparatively equal contribution of basal area, which highly correlated with biomass, from each sub-plot. We used the available *DBH* distribution data from earlier research projects that were collected in the study area by the team of Prof. Florian Siegert from RSS Gmbh, Germany, and the team of CIMTROP, Indonesia.

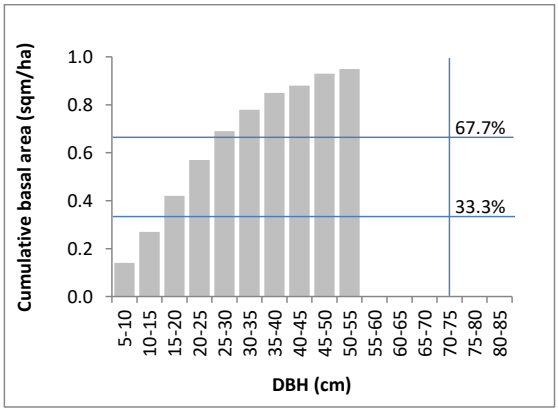
The questions of which trees to be measured in each sub plot, and how big the area of each sub-plot, have to define before the field campaign.

(i) Determination of the *DBH* thresholds in each sub-plot

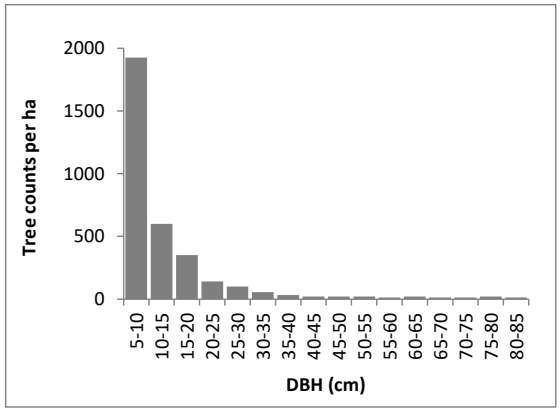
We estimated the tree basal area per hectare and calculated the cumulative basal area values. The *DBH* thresholds were determined by the proportional distribution of the cumulative basal area among sub-plots, which were three sub-plots in this study. We derived the *DBH* thresholds of approximately 17 cm for the smallest sub-plot and of approximately 30 cm for the medium sub-plots, from the cumulative basal area at 33.3% and at 66.7%, respectively. All *DBH* class above 30 cm or 66.7% of the cumulative basal area values belongs to the largest sub-plot (Figure 1).

(ii) Determination of the radius of each sub-plot

We estimated the number of trees per hectare over three *DBH* classes: 5 cm ≤ *DBH* < 17 cm, 17 cm ≤ *DBH* < 30 cm, and *DBH* ≥ 30 cm (Figure 2), and calculated the proportional area of having not less than 10 trees per each class. We derived the radius of each sub-plots from the size of the area of each class.

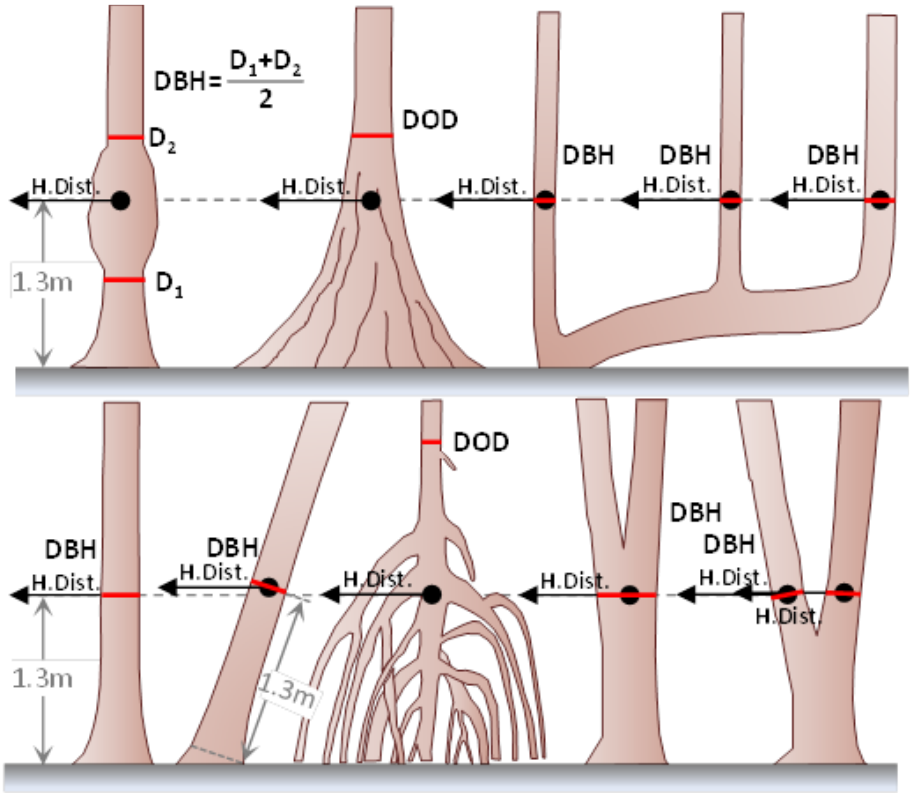


Proportional distribution of cumulative basal area over tree DBH



Trees per per hectare over DBH distribution

APPENDIX B. Definition of 1.3 m height of the tree *DBH* measurements



Definition of 1.3 m height used to measure the tree *DBH* in the field, for regular and irregular stems;

The upper figure, left to right: deformation tree, buttress tree, forking tree.

The lower figure, left to right: a straight tree with regular stem, leaning tree with regular stem, stilt-root tree, the bifurcate tree at a higher than 1.3 m, bifurcate tree at lower than 1.3 m.

APPENDIX C. List of the selected AGB models based on the EP Dataset 1 and 2

List of the selected AGB models based on Dataset 1.

Plot sides	Plot size	Selected predictor variables in the model				$Adj\ R^2$	$RMSE_r$
20x20	400	B4_max	B4_min	NDVI_min	B2_sd	0.33	27.70
25x20	500	CCCI_max	B2_min	B5_min	CCCI_sd	0.51	28.64
25x25	625	B2_mean	B3_mean	CCCI_max	NDVI_sd	0.53	20.76
33.3x20	667	B4_max	B5_max	NDVIRE_max	NDVIRE_min	0.49	22.33
33.3x25	833	B4_max	B5_min	CCCI_min	HOMO_X15	0.55	22.63
50x20	1000	B5_mean	B2_max	NDVIRE_max	NDRE_min	0.39	21.64
33.3x33.3	1111	B2_mean	B3_min	MSAVI_min	B2_sd	0.50	18.91
50x25	1250	B3_mean	B3_min	B4_min	CCCI_sd	0.49	19.70
50x33.3	1667	B2_mean	B5_max	NDVIRE_max	MEAN_X15	0.62	14.69
100x20	2000	B2_min	B4_min	B3_sd	B5_sd	0.55	14.16
100x25	2500	MSAVI_mean	B2_min	NDVI_min	NDRE_sd	0.45	13.97
50x50	2500	B3_mean	B4_max	B3_sd	MSAVI_sd	0.50	15.63
100x33.3	3333	NDVIRE_mean	B5_min	NDVI_max	CCCI_max	0.58	12.70
100x50	5000	CCCI_mean	MSAVI_max	NDRE_max	B5_min	0.51	12.78
100x100	10000	B3_mean	MSAVI_max	CCCI_min	NDRE_sd	0.65	10.27

List of the selected AGB models based on Dataset 2.

Plot sides	Plot size	Selected predictor variables in the model				$Adj\ R^2$	$RMSE_r$
20x20	400	B2_min	B2_sd	MEAN_X15	CORR_X15	0.51	23.05
25x25	625	B4_mean	B3_max	B3_sd	NDVI_mean	0.44	23.54
30x30	900	B4_max	B3_sd	B4_sd	MEAN_X15	0.54	16.36
35x35	1225	B2_min	B4_min	B3_sd	MEAN_X15	0.63	14.90
40x40	1600	B4_min	B3_sd	B5_sd	MEAN_X15	0.62	13.52
45x45	2025	NDRE_max	NDVI_max	B3_sd	B4_sd	0.59	14.06
50x50	2500	B3_mean	B3_max	NDVI_max	B5_min	0.56	14.56
55x55	3025	NDRE_mean	B3_max	B2_min	MEAN_X15	0.68	11.72
60x60	3600	B3_mean	B3_max	NDVIRE_max	B2_min	0.60	11.12
65x65	4225	B3_mean	B3_max	NDVIRE_max	B2_min	0.61	11.37
70x70	4900	NDRE_mean	B2_min	B3_sd	MEAN_X15	0.60	11.97
75x75	5625	MSAVI_max	NDRE_max	CCCI_min	MEAN_X15	0.49	15.07
80x80	6400	MSAVI_max	NDRE_max	B5_min	MEAN_X15	0.50	14.09
85x85	7225	MSAVI_max	NDRE_max	CCCI_min	MEAN_X15	0.54	12.94
90x90	8100	MSAVI_max	CCCI_min	NDRE_min	NDRE_sd	0.61	11.63
95x95	9025	MSAVI_max	NDRE_max	CCCI_min	MEAN_X15	0.53	12.33
100x100	10000	B3_mean	MSAVI_max	CCCI_min	NDRE_sd	0.65	10.27