Integrating Trees Outside Forests into National Forest Inventories

Sebastian Schnell
Faculty of Forest Sciences
Department of Forest Resource Management
Umeå

Doctoral Thesis
Swedish University of Agricultural Sciences
Umeå 2015

Acta Universitatis agriculturae Sueciae
2015:37

Cover: Single tree on sheep-run
(drawing by Edda Schnell)

ISSN 1652-6880
ISBN (print version) 978-91-576-8272-7
ISBN (electronic version) 978-91-576-8273-4

Print: SLU Service/Repro, Uppsala 2015

Integrating Trees Outside Forests into National Forest Inventories

Abstract

Trees Outside Forests (TOF) offer a wide range of ecological, economic, and social services. For example, they sequester carbon, provide wood for fuel and construction, protect soils from erosion, and contribute to the conservation of biological diversity. In particular in regions with low forest cover, TOF often have a substantial role in meeting society's demands for resources such as wood and fodder.

Information about trees is required for many purposes and at many geographical scales, and it has been recognised that substantial tree resources are overseen when focussing on forests alone. At the global scale, reporting obligations linked to agreements such as the Kyoto protocol are important. However, information is also needed for policy making at national scale and for integrated management by rural and urban planners. The focus of this thesis is the provision of national level information about TOF resources.

From a literature review it was concluded that many national forest inventories have widened the scope of their inventories through including TOF. However, in general there is a shortage of information about TOF resources on a global scale. Further, very few methodological studies exist on how TOF could be integrated into national forest inventories. A central question of this thesis thus is how an integrative monitoring approach such as a *national tree inventory* would look like.

Existing data from country-level TOF inventories across three continents were re-analysed. It was found that TOF contribute substantially to national tree biomass and carbon stocks. A method for simulating the spatial distribution of TOF elements at the landscape scale was investigated at selected study sites in Skåne, in the south of Sweden. The aim was to reconstruct existing patterns by methods from material sciences that might be used for modelling TOF patterns. Finally, a sampling simulation study was conducted to assess the potential of different inventory strategies to form the basis for national tree inventories. It was found that the combination of data from field sample plots and airborne laser scanning offers great potential in connection with model-assisted estimation.

The results of this thesis may serve as a starting point for moving from a forest-centred view on tree monitoring towards integrative monitoring approaches that consider all trees that grow in a study region as valuable.

Keywords: survey sampling, remote sensing, estimation, biomass

Author's address: Sebastian Schnell, SLU, Department of Forest Resource Management, Skogsmarksgränd, 90183, Umeå, Sweden
E-mail: Sebastian.Schnell@ slu.se

Dedication

To Edda, Ragnar and Randolf

„Habt Ehrfurcht vor dem Baum, er ist ein einziges großes Wunder, und euren Vorfahren war er heilig. Die Feindschaft gegen den Baum ist ein Zeichen von Minderwertigkeit eines Volkes und von niederer Gesinnung des einzelnen“

“Have reverence for the tree, he is a single great wonder, and was holy to your ancestors. The hostility against the tree is a sign of the inferiority of a nation and of an individual’s cavalier attitude”

Alexander von Humboldt

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Schnell, S., Kleinn, C, Ståhl, G. (2015). Monitoring trees outside forests: a review (submitted).

II Schnell, S., Altrell, D., Ståhl, G. and Kleinn, C. (2015). The contribution of trees outside forests to national tree biomass and carbon stocks – a comparative study across three continents. *Environmental Monitoring and Assessment* 187:4197 (1-18).

III Schnell, S. and Pommerening, A. (2015). Reconstructing tree cover patterns in the (open) landscape (manuscript).

IV Schnell, S., Ene, L.T., Grafström, A., Nord-Larsen, T. and Ståhl, G. (2015). Sampling strategies for large area inventories of trees outside forests - a simulation study (resubmitted).

Papers II is reproduced with the permission of the publisher.

The contributions of Sebastian Schnell to the papers included in this thesis were as follows:

I Planned the study with the co-authors. Carried out the literature review and wrote the major part of the manuscript.

II Planned the study with the co-authors. Carried out all data processing and analysis and wrote the major part of the manuscript.

III Planned the study with the co-author. Carried out all programming and data analysis and wrote the major part of the manuscript.

IV Planned the study with the co-authors. Carried out all programming and data analysis and wrote the major part of the manuscript.

Abbreviations

TOF	Trees Outside Forests
AGB	Aboveground Biomass
NFI	National Forest inventory
NFMA	National Forest Monitoring and Assessment
FAO	Food and Agricultural Organisation
F	Forest (FAO land use/cover class)
OWL	Other Wooded Land (FAO land use/cover class)
OL	Other Land (FAO land use/cover class)
IW	Inland Water (FAO land use/cover class)
IPCC	Intergovernmental Panel on Climate Change
LPM	Local Pivotal Method
SRSwoR	Simple Random Sampling without Replacement
PSU	Primary Sampling Units
SSU	Secondary Sampling Units
DBH	Diameter at Breast Height
IPCC	Intergovernmental Panel on Climate Change
DEM	Digital Elevation Model
ALS	Airborne Laser Scanning
LiDAR	Light Detection and Ranging
MODIS	Moderate Resolution Imaging Spectro-radiometer

1 Introduction

1.1 Background

Globally, substantial efforts are put into the monitoring of forest resources (e.g., FAO, 2010), and there are good reasons for this due to the multitude of ecosystem services provided by forests. However, tree resources that grow outside forests typically are not taken into account by forest monitoring programmes even though they provide similar services as forest trees. Such tree resources are summarised with the term trees outside forests (TOF). The term was coined by the Food and Agricultural Organisation (FAO) of the United Nations while planning the Global Forest Resources Assessment 2000 (FRA) in the mid-1990s (Pain-Orcet & Bellefontaine, 2004). By introducing the TOF concept, the FAO aimed at increasing political attention to TOF (de Foresta *et al.*, 2013) because it was recognised that considerable amounts of tree resources in many countries might be overlooked when monitoring focuses solely on forests. Furthermore, the relevance of TOF for human livelihood, general environmental conditions, and biodiversity was emphasised (Pain-Orcet & Bellefontaine, 2004; Bellefontaine *et al.*, 2002).

Interest in TOF has existed much longer than the recent attempts by the FAO to provide a more holistic view on tree resources. Already in 1713, Hans Carl von Carlowitz, in addition to introducing the term “sustainability” in the context of forest management, referred to non-forest tree resources several times in his opus *Silvicultura Oeconomica* (Carlowitz, 1713). For centuries, people have managed various agroforestry systems such as fruit tree meadows, hedgerows, riparian buffers, or parkland trees (Boffa, 2000; Herzog, 2000). In addition, there is a long history and wealth of literature from agroforestry and urban planning that emphasise the role of TOF. Also, TOF inventories have been implemented before (Holmgren *et al.*, 1994; Bieberstein *et al.*, 1982; Bieberstein *et al.*, 1975).

TOF are trees that grow on land not defined as forest or other wooded land (Bellefontaine *et al.*, 2002). This basic definition from FAO relies on the definitions of forest and other wooded land, and there are many forest definitions available that often vary between countries (Lund, 2002). Thus, the definition of what constitutes TOF can also vary between countries. However, at an international level, the FAO's definitions of forest and other wooded land are widely accepted and applied (FAO, 2010).

The forest definition of the FAO has a strong focus on land use, and it states that forest land must not be under predominantly agricultural or urban land use. In addition, a certain canopy cover, tree height, and area size is required for land to be considered a forest. By definition, this means that all trees on land with agricultural and urban use are considered TOF. In addition, tree formations with very low canopy cover or formations with a dense canopy but small areal extent are considered TOF regardless of land use. In addition, the word "trees" in TOF also includes shrubs, palms, and bamboo, which are relevant resources in some countries (de Foresta *et al.*, 2013).

As indicated by the definition, the TOF concept comprises a large variety of tree formations of various functional types and spatial arrangements. Classifications are difficult because many criteria exist for generating meaningful categories (Kleinn, 2000). Examples of such criteria are the land use on which TOF grow, the geometric patterns formed by the trees, the functions of the trees, and the origin of the trees.

Based on land use, TOF can be classified as growing on built-up areas or as growing on agricultural lands (Bellefontaine *et al.*, 2002). In built-up areas, possible formations include gardens, urban parks, street trees, hedgerows, trees along lakes and rivers, or production systems (e.g., orchards). On agricultural land, trees grow in combination with annual crops (coffee, cocoa, and other agroforestry systems), in pastures and meadows (scattered trees, windbreaks, woodlots, and other silvopastoral systems), along watercourses and bodies, or in orchards and permanent crop systems. Apart from urban and agricultural lands, TOF can also grow on natural lands as part of savannahs, at the tree line in mountainous areas, or on peat lands. Tree density in TOF is, therefore, typically very low.

The spatial arrangement of TOF can be categorized as isolated and scattered trees, stand-like groups of trees (e.g., small woodlots, parks), and plantings along linear landscape features (Alexandre *et al.*, 1999). This spatial diversity needs to be taken into account in monitoring systems by adjusting the inventory design to efficiently differentiate among the categories (Kleinn, 2000).

Functional classification would include TOF for the production of food, fodder, and firewood; TOF for the protection of agricultural crops and animals; and TOF for the provision of scenic beauty and other ornamental values. Applying origin as a classification criterion, the distinction is made between planted trees and trees that are remnants from former forests. The latter case is especially common in Latin America as a result of harvesting virgin forests (Kleinn, 1999).

A clear distinction from a true forest can be difficult because some ambiguities in the definitions exist. Kleinn (1999), for example, highlighted the following problem areas where clear distinctions from forest are difficult to draw: savannahs, coffee and cocoa plantations under shade trees, pasture lands with trees of varying densities, and orchards. These problem areas are mainly relevant for supporting inventories with remote sensing data because land use cannot be readily assessed and TOF formations might resemble forests on satellite images.

Other ambiguities, also related to definitions, were identified by de Foresta *et al.* (2013). Problematic areas include shifting cultivation, rubber plantations, linear tree formations, certain agroforestry practices, and agricultural or urban land use.

Shifting cultivation is a land-use form where short agricultural uses and longer periods, where the land lies fallow, alternate in a temporal sequence. During a fallow period, secondary forest re-grows and it is often unclear whether or not the land is forest and the agricultural land use has been abandoned. Trees on such land can thus be falsely counted as either forest or TOF.

Similarly, it is debated if rubber plantations should be counted as forest or TOF. The primary product is rubber, and timber is only a secondary product, available at the end of the rotation cycle. Currently, rubber plantations are seen as forest but were previously regarded as an agricultural cash crop (de Foresta *et al.*, 2013).

The assignment of linear tree formations to either forest or TOF can be problematic because land use, as well as geometric criteria, needs to be considered. Following the FAO’s forest definition, these tree formations are considered a forest when they are wider than 20 m and larger than 0.5 ha and when the land use is non-agricultural or non-urban.

Agroforestry systems are mostly considered TOF because of the predominant agricultural land use; however, cases exist where a forest land use dominates. For example, temporary grazing in forest plantations or the intercropping in the first years after plantation establishment would be considered as systems where the forest component dominates.

The meaning and exact definition of agricultural and urban lands differ between countries (de Foresta *et al.*, 2013). Also, the term "predominantly" is not further specified in the FAO definition, and this leaves room for interpretations. For example, the term urban forest does not mean that the land use is predominantly urban. In fact, the term has a relatively wide meaning and refers to all forest and tree resources within and close to urban areas (Konijnendijk, 2003). This means that urban forests are partly included in forest inventories and partly not. To separate urban forest from urban TOF, Rydberg and Falck (2000) suggest using the ground vegetation as a decision criterion; if uncultivated, it is an urban forest, otherwise it belongs to the realm of TOF.

From a climate change perspective, TOF are seen as a mitigation strategy because additional tree plantings on agricultural and urban lands for carbon sequestration normally do not compete with other land uses (Schoeneberger *et al.*, 2012; Plieninger, 2011; Schoeneberger, 2009). The potential for carbon sequestration through TOF can thereby be considered to be high, in particular for agroforestry systems with a rather dense canopy cover (IPCC, 2000, Table 4-1). In addition, important co-benefits, such as erosion control (Manning *et al.*, 2009; Baudry *et al.*, 2000) and the conservation and improvement of biological diversity (Paletto & Chincarini, 2012; Bhagwat *et al.*, 2008) can be obtained. TOF are also used as a source for food (Herzog, 1998), fodder (Hinsley & Bellamy, 2000), and wood products (Ahmed, 2008; Pandey, 2008). For many societies, the cultural value (Grala *et al.*, 2010) and scenic or recreational uses are of importance (Herzog, 2000). Even in urban environments carbon sequestration can be substantial (Nowak, 2002) in addition to other services like regulation of micro-climate (Bowler *et al.*, 2010) and removal of air pollutants (Jim & Chen, 2009).

1.2 Motivation for monitoring TOF

From the previous section it can be concluded that TOF provide many goods and services to society. Especially in developing countries with sparse forest cover, TOF may be a major source of wood resources. Information about TOF is needed at several geographical scales. At the international scale, agreements such as the United Nations Framework Convention on Climate Change (UNFCCC, 2008) and its Kyoto protocol demand information about all tree resources, not only trees in forests. In this case, the focus is biomass and biomass change as proxies for carbon dioxide emissions and removals. Examples of international agreements that require information of more general kind about TOF is the United Nations Convention of Biological Diversity

(SCBD, 2005) and the United Nations Convention to Combat Desertification (UNCCD, 1994).

At national scale there is a general need for information to support policies and legislation related to the conservation and use of TOF resources. In this case TOF information would have a potential to provide input to several sectorial fields, such as agriculture and urban planning. In addition, the information reported under the international agreements normally is acquired and at the national scale.

At local scale there is a need to acquire information to support the management of TOF. Such information would need to be adapted to the specific management goals, which would vary substantially depending on what subcategory of TOF is considered. For example, MacFarlane (2009) concludes that biomass derived from urban trees offers a potential (1) for deriving local wood products, (2) for locally generated fuel sources to generate power and heat, (3) for reducing fossil fuel consumption, (4) for reducing waste disposal costs, and (5) for reducing pressure on forests.

1.3 Integration into national forest inventories (NFIs)

As a consequence of the FAO's attempts to encourage the monitoring of TOF, a number of countries have widened the scope of their NFIs to include TOF. Examples are the NFIs of India (Tewari *et al.*, 2014), Sweden (Fridman *et al.*, 2014), France (Bélouard & Coulon, 2002), and Switzerland (Brändli, 2010). A more complete list can be found in de Foresta *et al.* (2013). In addition all countries that conduct their NFIs under the National Forest Monitoring and Assessment (NFMA) programme of the FAO include TOF by default (FAO, 2012). Also, several pilot studies within the framework of the forest inventory and analysis programme of the USA have been conducted (Lister *et al.*, 2012; Cumming *et al.*, 2008; Riemann, 2003).

The NFIs listed above show that substantial wood resources can be overlooked if TOF are not included in the monitoring systems. For example, in India it was reported that 25.6 % of the national growing stock of trees exists outside forests (FSI, 2011). For the states Haryana, Kerala, and Punjab in India, the majority of wood is actually produced by TOF (Ahmed, 2008; Pandey, 2008). For the NFMA inventories, the share of TOF to national tree aboveground biomass (AGB) stocks varied between 3 % and 70 % (Schnell *et al.*, 2015). Lister *et al.* (2012) estimated that 19 % of all trees in the Great Plains states of the USA grow outside forests. A similar proportion was found by Riemann (2003) for five counties in Maryland, USA. For forest-rich countries like Sweden, however, the contribution of TOF to national growing stocks is very low (0.7%).

Urban areas often have a relatively high tree cover, highlighting the importance of TOF in such environments. For example, for urban areas in Wisconsin it was estimated that 26.7 % of the area was covered by trees (Cumming *et al.*, 2007). Further, Nowak (2002) found that average carbon stocks ranged between 4.4 Mg ha^{-1} and 36.1 Mg ha^{-1} for ten cities in the USA. In another study, Nowak *et al.* (2008) found that tree cover for 14 US cities ranged from 8.9 % to 36.7 %.

Despite the general acceptance of the importance of TOF and advances in monitoring, data that would be needed for an integrated management of landscapes for climate change mitigation and adaptation (Plieninger, 2011) is, in general, still missing at the global scale and only partly available at the national scale (de Foresta *et al.*, 2013). One reason for this is that not all kinds of TOF are included in the monitoring. For example, in Sweden TOF is generally included in the NFI but with the exception of trees growing in human settlements, thus only allowing conclusions to be drawn about a specific subset of TOF. Another reason is that even though assessments are done in many countries, results for TOF are hardly ever reported publicly and are difficult to access.

A further problematic area for TOF inventories is the involvement of many stakeholders. TOF occur across the agricultural, the forest, and the urban sectors. When monitoring crops and livestock, the agricultural sector ignores trees, and, likewise, the forest sector concentrates mainly on forest trees (de Foresta *et al.*, 2013). Thus, a potentially large tree resource, as shown above, is often overlooked. In addition, forest inventories often have no mandate to include TOF (Perry *et al.*, 2009). However, because of the rich experience in monitoring trees over large areas and because of the existing infrastructure, it could be argued that TOF should be integrated into NFIs.

2 Objectives

The main objectives of this thesis were to investigate the importance of TOF in different countries, to assess the current state of the art regarding TOF monitoring, and to evaluate methodological options for integrating TOF into NFIs or similar inventories. The focus was set on inventories that provide national level information, including information for international reporting. Thus, the studies deal with general properties of TOF, such as abundance patterns and biomass assessment.

The specific objectives of Papers I–IV were:

- I To provide a general overview of TOF inventory methodologies through a literature review. The intention was to assess the general status of TOF monitoring, identify good examples and interesting methodological opportunities, and suggest important areas for future study.
- II To demonstrate the importance of TOF resources in different parts of the world through re-analysing existing data from different countries using standardised methodology, thus facilitating detailed comparisons. Further, the sensitivity of the results to the application of different modelling approaches was analysed.
- III To develop methods for modelling and reconstructing spatial patterns of TOF resources based on empirical landscape data. The methods can be applied for testing the suitability of functional summary statistics for describing different features of TOF landscapes and for providing fictitious landscapes for sampling simulation.
- IV To compare the efficiency of different large area sampling strategies for integrating monitoring of TOF into NFIs. The study was conducted through sampling simulation.

Due to space limitations, only a basic set of results from the compilation of TOF resources in different countries was presented in Paper II. Thus, some additional results from the study are included in this thesis summary to provide a more complete picture.

The original ambition was to utilise TOF landscapes created with the method proposed in Paper III for the sampling simulations in Paper IV. However, during the course of the work it was found that a slightly different approach was more appropriate for the purpose of Paper IV.

3 Material and Methods

3.1 Methodological approach of the thesis

No formal hypotheses were formulated and tested as part of the studies, as the work mainly involved reviews and the development and evaluation of methods. This corresponds to the normal practice for scientific studies in the fields of survey sampling and forest inventory methodology, where the performances of different methods are either compared strictly analytically or through quantitative approaches where properties, such as the precision of a sampling strategy, are assessed and compared numerically without prior assumptions about which method should perform best.

3.2 Reviewing literature on existing inventory methodologies

A literature review was conducted in order to assess the current state of the art regarding inventories of TOF. The review focused on selected aspects of relevance for implementing large area monitoring of TOF. These were (1) TOF definitions; (2) field-based inventory systems; (3) what sampling units have been used in different studies; (4) how remote sensing data might contribute to the monitoring; (5) strategies for combining field-based and remotely sensed data; and (6) the availability of allometric models for estimating the biomass of individual trees.

The literature was searched through the help of keywords in the ISI Web of Knowledge. As keywords, terms that are frequently used as synonyms for TOF or that describe a specific TOF subset were used. Examples are non-forest trees, working trees, scattered trees, urban trees, hedgerow, shelterbelt, etc. In addition terms describing the action of monitoring like inventory, sampling, survey, etc. were used to filter the search results.

Query results were screened based on titles, abstracts, and author keywords, and articles were chosen subjectively for their relevance with respect to the aspects listed above. Articles published before 1990 were not actively searched for because the interest in monitoring TOF at national scale essentially started around 1995, when FAO decided to include TOF into their global forest resources assessments (Pain-Orcet & Bellefontaine, 2004). Another reason for the limit was to focus only on the more recent developments in the field of general methods for natural resources inventories.

3.3 Reanalysing existing inventory data towards TOF

In Paper II, existing data from the NFMA programme of the FAO were reanalysed to provide comparative information about the importance of TOF resources in different regions of the world. Such a compilation did not exist prior to this study.

3.3.1 Data

In July 2011, data for 11 countries that had completed their inventories under the NFMA programme were made available for the study. The 11 countries are located on three different continents and mainly in the tropics (*Figure 1*). An overview of general country characteristics is given in *Table 1*. Land areas range from about 10,000 km² (Lebanon) to as much as 750,000 km² (Zambia) with a great variety in population density, which is highest in Bangladesh with about 1,000 people/km². Also, the Philippines and Lebanon have relatively high population densities compared to the other countries, which all have less than 100 people/km². The economic conditions vary considerably with the average annual gross domestic product per capita ranging from less than 1,000 US$ to about 10,000 US$.

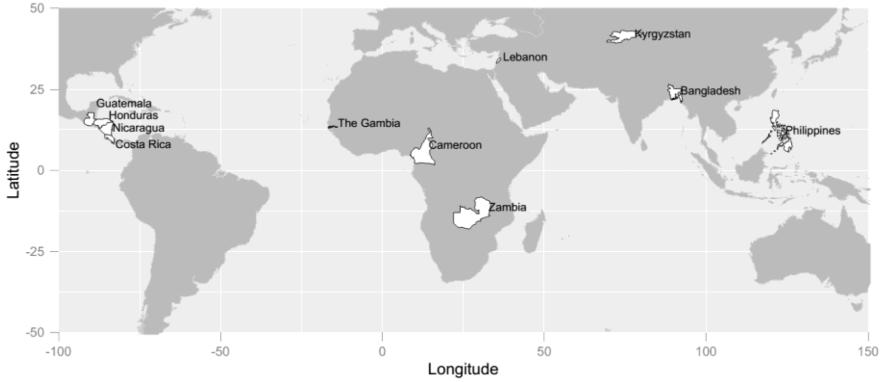

Figure 1. Map of countries that are included in the analysis in Paper II.

3.3.2 Analysing the data

The NFMA inventory data were re-analysed using the estimation framework for a mapped plot design described in Zarnoch and Bechtold (2000). In the NFMA inventories, clusters of field plots were distributed systematically over the study area. Each cluster consisted of four rectangular sub-plots with a size of 20 m × 250 m, so that a plot area of 2.0 ha was tallied at each sample location (*Figure 2*). The country's whole territory was used as the sampling frame, and all cluster plots were visited in the field. Trees were measured regardless of the land use, i.e. the plot area was mapped according to actual land use. The three global FAO land-use categories forest (F), other wooded land (OWL), and other land (OL) were used for the break-down of results. Trees growing in OL are considered as TOF according to the definition by FAO (de Foresta *et al.*, 2013).

Because countries can adjust the basic NFMA inventory design to local conditions, an overview of design characteristics is given in *Table 2*. It should be noted that Kyrgyzstan and Zambia used smaller DBH thresholds in order to include trees below the standard threshold of 10 cm. These small variations between countries were not corrected for, and results should be interpreted with that in mind. Also the sample sizes (the number of cluster plots) varied considerably between the studied countries, largely as a result of variability in country size and available financial resources.

For each tree and for each land use encountered on the cluster plots, several variables were observed. For this thesis, only observations of DBH, height, and tree species were used, as well as the area and the land-use category of the plot sections on which the trees were found. For estimation, tree-based observations were aggregated to the cluster level and broken down by the major land-use categories.

Table 1. *Population, gross domestic product, and climate for the eleven countries studied in Paper I (World Bank, 2013; Mitchell et al., 2004).*

Country	Population		GDP (US$, nominal, 2011)		Temperature[3] (1961–1990)	Rainfall[4] (1961–1990)
	Total[1]	per km^2	Total[2]	per capita	℃	mm
Bangladesh	150.49	1020	111.88	743	25.0 (18.5–27.8)	2666 (7–596)
Cameroon	20.03	42	25.24	1260	24.5 (23.3–26.5)	1603 (14–265)
Costa Rica	4.73	92	40.87	8647	24.8 (23.9–26.1)	2926 (73–413)
The Gambia	1.78	171	0.90	506	27.5 (24.1–30.5)	837 (0–267)
Guatemala	14.76	136	46.90	3178	23.4 (21.1–25.2)	2713 (66–420)
Honduras	7.75	69	17.43	2247	23.5 (21.3–25.1)	1976 (42–263)
Kyrgyzstan	5.51	28	6.20	1124	1.6 (–13.8–14.9)	381 (12–58)
Lebanon	4.26	408	40.09	9413	16.4 (7.7–24.8)	661 (0–140)
Nicaragua	5.87	45	9.32	1587	24.8 (23.5–26.1)	2391 (41–351)
Philippines	94.85	316	224.77	2370	25.8 (22.5–27.0)	2348 (85–277)
Zambia	13.47	18	19.21	1425	21.4 (16.9–24.4)	1020 (0–229)

1 Million inhabitants
2 In billions
3 Yearly average temperatures with monthly averages for the coolest and hottest month in parentheses
4 Yearly average precipitation with the precipitation of the driest and wettest month in parentheses

Because allometric models that estimate AGB at the individual tree level are extremely rare for TOF (McHale *et al.*, 2009), mixed-species pan-tropical biomass equations were used (Chave *et al.*, 2005). This is in agreement with suggestions from Nair (2012) for estimating AGB of agroforestry systems, and this method is also commonly applied in NFIs in the tropics (Chave *et al.*, 2014).

For the two non-tropical countries, Kyrgyzstan and Lebanon, biomass estimations were based on the inventoried volume, that was converted to AGB using average wood specific gravity and biomass expansion

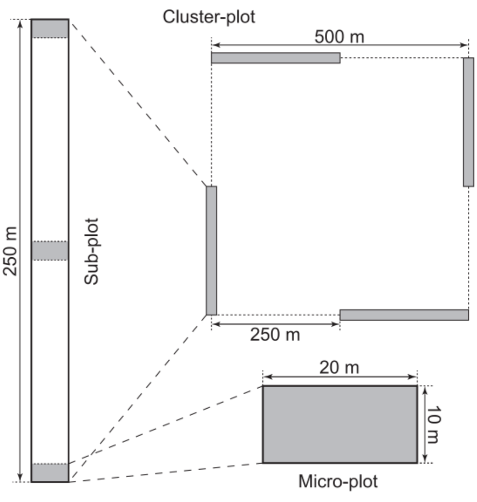

Figure 2. Basic plot design used in the NFMA programme. Only the elements necessary for the analysis in Paper II are illustrated. A detailed description is given by the FAO (2012). The application of micro-plots was optional and not applied equally by all countries. For Cost Rica, the subplot length was 150 m.

factors from IPCC tables (IPCC, 2006). This is in agreement with the methodology that was applied by the countries for their own reporting of NFMA results. To estimate tree volume, species group-specific form factors were applied for each species.

Table 2. *Basic characteristics of the sample and plot design used in the different countries (Paper I). Sample size is the number of clusters used; in case of inaccessibility the reduced number of clusters is given in parentheses. The diameter at breast height (DBH) limit is the threshold for measuring trees on forest sub-plots, with thresholds for trees on rectangular nested plots and outside forests given in parentheses. The binary variable Nested plot indicates whether nested plots were used (1) or not (0).*

Country	Area	Sample size	Strata	DBH limit	Nested plot	Year
	km^2	no.	no.	cm	bin	
Bangladesh	147 570	298	1	10 (10)	0	2005–2007
Cameroon	475 440	207 (205)	2	20 (10)	1	2003–2004
Costa Rica	51 000	40 (39)	1	30 (10)	1	2001
The Gambia	11 000	144 (129)	1	20 (10)	1	2008–2010
Guatemala	108 899	114	3	20 (10)	1	2002–2003
Honduras	112 492	181	1	20 (10)	1	2005–2006
Kyrgyzstan	199 940	765 (733)	1	8 (8)	0	2008–2010
Lebanon	10 452	222 (220)	1	10 (10)	0	2003–2005
Nicaragua	130 000	371 (368)	1	20 (10)	1	2007–2008
Philippines	300 000	351 (349)	1	20 (10)	1	2003–2005
Zambia	752 614	238 (232)	1	20 (7)	1	2005–2008

The biomass of palms was estimated with the same models as those used for trees, despite the fundamentally different allometry and highly variable wood density within single palm trunks (Rich, 1987). The reason for this was mainly the lack of allometric equations for biomass estimation of palm trees based on the input variables (diameter and height) that were available from the inventory data. Models based on trunk height exist for some Mesoamerican species (e.g., Saldarriaga *et al.*, 1988; Frangi & Lugo, 1985). However, these palm species differ substantially in average size and height, especially compared to Asian species that are, e.g., frequent in Bangladesh or in the Philippines. The lack of allometric biomass models for palms appears to be a major methodological gap for TOF and forest biomass estimation in tropical regions.

3.4 Reconstructing tree cover patterns

The initial idea for Paper III was to simulate maps of TOF cover that could serve as an input for simulation studies of sampling strategies (Paper IV). The methodological approach was to use techniques from material physics (e.g., Yeong & Torquato, 1998) to reconstruct existing TOF patterns as extracted from remote sensing data. The reconstructions were based on simulated annealing, an optimisation algorithm introduced by Kirkpatrick *et al.* (1983). The morphological information contained in original patterns is described using functional summary statistics. The same summary statistics are applied to random configurations that are changed at each simulation step. If the random configuration improves, i.e. if the summary statistics of the original and random configuration become more similar, the original pattern is updated. The procedure requires a large number of iterations and proceeds step by step towards the target configuration. The procedure depends on how well the summary statistics reflect the morphological properties of the original pattern.

Using this technique, the original intention was to identify suitable statistical summary functions for describing patterns of TOF cover as extracted from remote sensing data. After that the idea was to identify mathematical models that could replace the statistical summary functions. From these models it was thought to generate different spatial configurations of TOF in order to investigate the effect of spatial characteristics on different sampling strategies. However, during the course of the work it turned out that the original idea could not be implemented, partly because the simulated annealing technique only worked for a limited spatial extent (in the order of 10 hectares) due to computational limits.

The reconstruction algorithm is generally applied to regular lattices, consisting of single cells, typically representing either pores or the skeletal structure of materials. However, there are also applications to spatial point patterns (Tscheschel & Stoyan, 2006). For forest ecosystems, two main types of application can be distinguished: (1) reconstruction of the patterns of tree locations in forest stands from complete observations (Pommerening, 2006); and (2) reconstruction of the same patterns from partial or incomplete information, e.g. from forest inventory data (Pommerening & Stoyan, 2008). Further, the technique allows for studies of the underlying point processes and the assessment of habitat suitability for meeting requirements of nature conservation (Bäuerle & Nothdurft, 2011), or to locally reconstruct tree locations around inventory plots for improving sampling estimators (Lilleleht *et al.*, 2013; Nothdurft *et al.*, 2010).

In the following, in deviation from the original plan, a first attempt to apply stochastic reconstruction to patterns of tree cover is presented. The tree cover patterns outside forests differ from many other random heterogeneous media since the phase of interest (canopy cover) is comparably rare compared to the background and may have a pronounced non-random spatial distribution (e.g., linear features).

3.4.1 Study area and maps of canopy cover

The study area was the county of Skåne in the south of Sweden, which has a, relatively low forest cover (35 %) compared to other parts of Sweden (*Figure 3*). For identifying tree vegetation, airborne laser scanning (ALS) data were used; it was acquired from a national flight campaign conducted for deriving a new digital elevation model (DEM) for Sweden. In order to process the ALS data and to extract tree canopy cover, additional data in the form of a DEM and land-use information were required. The DEM was provided by the Swedish Land Survey Agency (Lantmäteriet, 2014b) and had a pixel size of 2 m × 2 m.

For the land-use information, the estate map and the general map of Sweden (Lantmäteriet, 2014a; Lantmäteriet, 2014c) were used. Layers describing building outlines, human settlements, and a forest mask were extracted. The forest mask was cleaned from polygons smaller than 0.5 ha and narrower than 20 m in order to fulfil the FAO forest definition (de Foresta *et al.*, 2013, p. 28). In addition, for the parts of the forest polygons that intersected with settlements, a non-forest land use was assumed and tree cover in such areas was treated as TOF.

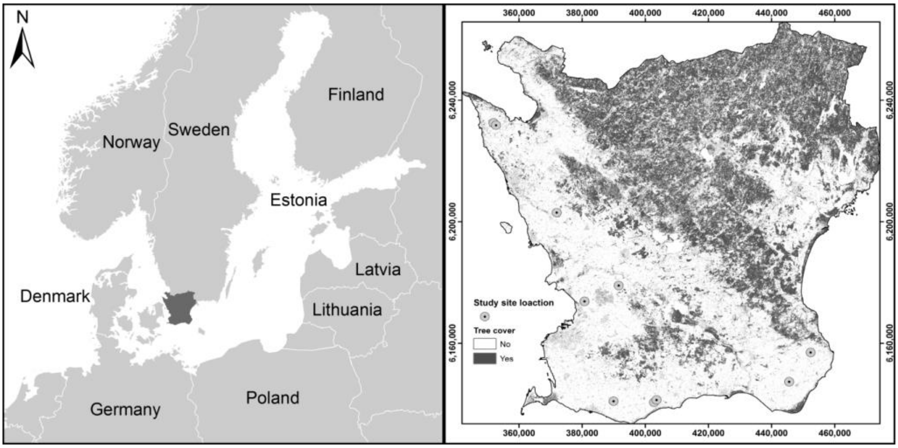

Figure 3. Location of the selected study site: Skåne, Sweden (area shaded in dark grey on the left panel). The dots indicate the locations of the lower left corners of the 400 m × 400 m large images that were used for the reconstruction (right panel).

The ALS point cloud was normalised to match the DEM in order to obtain height above ground information. The building mask was used to classify the point cloud into points representing buildings and other points. To identify vegetation, a local transparency and roughness measure – the slope adaptive echo ratio – was calculated for each point (Höfle *et al.*, 2009). The echo ratio of single points was then averaged over a 2 m × 2 m raster excluding the points that were previously classified as buildings. Cells that had an average echo ratio of less than 85 % were classified as vegetation. The preliminary vegetation mask was refined by removing any remaining artificial objects such as remnants of buildings, power lines, street lights, etc. by the technique described in Eysn *et al.* (2012). Finally, information from the land-use maps was combined with the final tree canopy mask to yield three tree cover categories: forest trees, trees in settlements, and trees in the open landscape.

To select specific TOF patterns, the tree cover map (*Figure 3*) was divided into 200 m × 200 m squares. The focus was on areas that were completely located in Skåne and that had no trees in forests or settlements. In total, 1,625 squares fulfilled these criteria. From this set of squares, 10 were selected by the sampling protocol described in the following. FRAGSTAT (McGarigal *et al.*, 2012) was used to calculate a set of metrics that described the general abundance and shape of TOF patches in the single squares. The following metrics were chosen: percentage of area covered by patches, largest patch index, average patch size, average of the related circumscribing circle, and the perimeter-area fractal dimension. With the help of these metrics, a sample of 10 squares was chosen using the local pivotal method (Grafström *et al*. 2012). With this sampling technique, auxiliary variables are used to select samples that are evenly distributed in the space defined by the auxiliaries. The result was a sample of images that represents the variability of the available TOF patterns with respect to the used auxiliaries. The selected images constituted the original images for the reconstruction experiment.

3.4.2 Summary statistics

To reconstruct the original images, summary statistics that describe the spatial properties of the TOF patterns in the original and reconstructed images were needed. The two-point probability function and the lineal-path function were used for this purpose (Gerke *et al.*, 2014; Yeong & Torquato, 1998).

The two-point probability function $S_2(x_1, x_2)$ is a measure of the probability that the two points x_1 and x_2 are both covered by the same image category, i.e. tree cover in the present case. The lineal-path function $L(\mathrm{x}_1, \mathrm{x}_2)$ also uses two points x_1 and x_2. Here, however, the function assesses whether the line spanned by the two points is completely covered by trees.

The two functions describe different aspects of tree-cover patterns (Yeong & Torquato, 1998). The connectivity of tree pixels is covered by the lineal-path function, but it does not contain any information beyond the maximum cluster size. This short-range information of different clusters is, however, addressed by S_2; thus the two functions complement each other. Another reason for selecting these two functions is their computational nature. They are comparatively easy to implement and can be calculated quickly, which is important for the reconstruction process.

When calculating L and S_2, all possible point pairs in an image are normally evaluated and summarised by the distances between the point pairs. The final functions thus represent the frequency distribution of successful events, i.e. both points represent trees or a straight line connecting two points is covered by trees. Additional morphological information was gathered by evaluating the functions only along the two orthogonal and the two diagonal axes of the images (Gerke *et al.*, 2014; Yeong & Torquato, 1998).

3.4.3 Reconstruction method

The TOF cover patterns were reconstructed using simulated annealing with an improvements-only algorithm (Tscheschel & Stoyan, 2006). With this approach, a random start configuration or image that has the same number of tree-cover pixels as the respective original image was first generated. This random image will gradually evolve into a reconstructed image that mimics the original image in terms of the applied summary statistics L and S_2. To control the reconstruction of the original, a so-called energy or contrast measure E is used, which basically measures the squared differences between two functions (Yeong & Torquato, 1998).

In each simulation step, m, one background pixel and one tree pixel were randomly chosen and their states interchanged, i.e. background to tree and tree to background. By this a tree pixel was moved to a new random location. The summary statistics of the simulated pattern were updated after each simulation step, and a new energy $E_{\text{new}}$was calculated. E_{new} was then compared to the energy of the former step E_{old}. If $E_{\text{new}} < E_{\text{old}}$, the new configuration was accepted. If not, the previous configuration was restored. Thus, only improvements (lower energy) were accepted at each simulation step. The simulation was stopped when E reached a small value (E_{min}) or after 1×10^6 simulation steps. The implementation of the simulated annealing algorithm is summarised in *Figure 4* below.

The algorithm was programmed in C++ and run in an R script (R Core Team, 2014) using the `Rcpp` package (Eddelbuettel, 2013). During the simulation, periodic boundary conditions were applied to the images as a method to correct for edge effects (Illian *et al.*, 2008, p. 184; Yeong & Torquato, 1998). The simulation was completed after approximately four days.

Step 1: Initialisation

- Load original image
- Generate random image
- Initialise L^d and S_2 for both images
- Calculate E_{old}

Step 2: Move a pixel

- Randomly select one black and one white pixel
- Swap their states (black/white)
- Update L^d and S_2
- Calculate E_{new}

Step 3: Acceptance/rejection of the new configuration

- IF $E_{new} < \mathrm{E}_{old}$ accept and set $E_{old} = E_{new}$
- ELSE restore the old configuration and return to Step 2

Step 4: Quit the loop

- IF $E < \mathrm{E}_{min}$ OR $m < 1\times10^{-6}$ go to Step 2
- ELSE STOP

Figure 4. Pseudo code describing the simulated annealing algorithm (improvements-only).

3.5 Simulation of sampling strategies

3.5.1 Study area

For this paper, the same study area (Skåne) as for Paper III was used because of the low forest cover and thus the availability of larger amounts of TOF. Instead of selecting a few rather small study sites, the entire county was used as the sampling frame for the investigated sampling strategies. The distribution of tree AGBs for Skåne as derived from this study is shown in *Figure 5*. The same ALS data as used for Paper III were used, but for purposes of (1) generalising the field data to build the artificial population, (2) building an ALS AGB model that was necessary for some of the investigated sampling strategies, and (3) providing the auxiliary information used in some of the investigated sampling strategies.

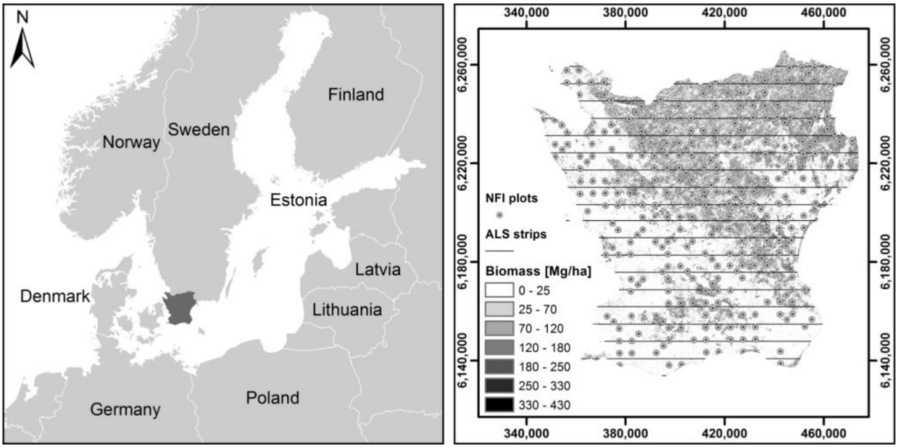

Figure 5. Study site location. Left: the location of Skåne in southern Sweden (area shaded in dark grey). Right: Aboveground tree biomass and the spatial location of permanent NFI clusters. Each point represents one cluster of at most four field plots. The horizontal lines indicate the location of 500 m wide strips that run east-west across Skåne. The strips are used in a two-phase sampling strategy for estimating total biomass. A selection of 20 out of 265 strips is shown in the figure.

3.5.2 Field data

Field data for building the artificial population and for developing the ALS AGB model were available from the Swedish NFI (Fridman et al., 2014). Permanent plots from the years 2008 to 2012 were used and a total of 1,033 plots were available. For each plot, AGB in Mg ha^{-1} was calculated using the models from Marklund (1988) for Norway spruce, Scots pine, and birch.

Field plots were available from all land-use categories except built-up areas. The majority of plots were situated on agricultural fields followed by forests and pastures (see *Table 3*). Most of the agricultural and pasture plots contained no tree biomass at all and thus received an AGB of zero. The plots were separated into a forest and a non-forest set, and no further distinctions were made between OWL and OL, even though the FAO makes such distinctions in its land-use definitions (de Foresta et al., 2013). However, in Skåne, only 5,000 ha are covered by OWL, thus justifying this simplification in building the artificial population. Non-forest land uses include fields, pastures, mires, roads, and power lines. Trees growing in settlements were not observed in the field but were all the same included in the final artificial population, which is explained further below.

Table 3. *Summary of the field data. Given are the total number of permanent plots available and the number of permanent plots that had some tree vegetation (non-zero). The statistics under 'Biomass' are based on the non-zero plots. Under* ***sd*** *the standard deviation and under* ***skew*** *the skewness of the data is given.*

Land use	Permanent plots		AGB (Mg ha^{-1})				
	total	non-zero	mean	min	max	SD	Skew
Field	432	11	45.6	3.4	144.7	39	1.2
Forest	417	398	109.3	0.4	429.4	72.4	1
Pasture	55	25	38.9	0.9	141.9	32.8	1.3
Mire	14	13	40.6	2.4	85.9	27.3	0.2
Road	6	2	44.6	9.6	79.6	49.5	0
Bedrock	3	0	0.0	0.0	0.0	n.a.	n.a.
Power line	2	2	21.7	0.9	42.5	29.4	0
Other	2	0	0.0	0.0	0.0	n.a.	n.a.
All	931	451	101.2	0.4	429.4	72.4	1.8

3.5.3 Building the artificial population

The vegetation mask and the map data from Paper III were used to classify the ALS point cloud into vegetation echoes, building echoes, and echoes coming from other land cover (mainly bare soil). The vegetation echoes were tessellated into 20 m × 20 m square cells covering the entire area of Skåne. For each square, a set of ALS metrics was calculated following an area-based approach (e.g., Næsset, 2002). The average height of the vegetation echoes (h_{mean}) and the ratio between vegetation echoes and the total number of echoes (p_{veg}) as a density measure were chosen as ALS metrics. The same metrics were calculated from the ALS points within the boundaries of the NFI field plots. Building echoes were excluded from these calculations.

The field-plot metrics together with the field-assessed AGB served as reference data. However, instead of linking the reference data directly to the full-cover ALS data via nearest-neighbour imputation, a copula model was introduced as an intermediate step as suggested by Ene *et al.* (2012) to avoid too many duplicated observations in the final product. From the fitted copula model, a large set of 250,000 observations of agb, h_{mean}, and p_{veg} was generated and merged with the original field data. First-nearest-neighbour imputation was then used to link the copula data to the full-cover ALS data by searching the first nearest neighbour in the space of the ALS metrics. The Euclidean distance was used for all measurements.

Based on the map data (forest mask and settlements), grid cells were assigned to one of the following domains: forest (F), settlements (SM), and other open land (OL, mainly agricultural fields and pastures).

3.5.4 ALS biomass model

For all of the investigated sampling strategies, model-assisted estimation (Särndal *et al.*, 1992) was applied as an alternative estimation technique. For this, a model was required that predicts the target variable agb from the ALS data. A model of the form

$$agb = \beta_0 \, h_{\text{mean}}^{\beta_1} \, p_{\text{veg}}^{\beta_2} \, \epsilon \qquad [3.5.1]$$

was fitted using ordinary least-squares regressions after logarithmic transformation of the response and predictor variables. Back-transformation bias was corrected using the ratio-like correction factor proposed by Snowdon (1991) and defined as $\sum agb / \sum a\hat{g}b$. Because dummy variables for land use had no significant effect, a common model was applied for all land-use categories. The leave-one-out cross-validated root mean square error (RMSE) of this model was 39.7 Mg ha^{-1} (38.5 %). The model performed much better for forest than for non-forest areas, and the cross-validated RMSEs for the forest and TOF domains were 41.1 Mg ha^{-1} (37.1 %) and 29.8 Mg ha^{-1} (68.8 %), respectively.

3.5.5 Local pivotal method (LPM)

The LPM was used as an alternative to simple random sampling with replacement (SRSwoR) to evaluate the effect of using auxiliary information for the selection of samples. The LPM was first introduced by Grafström *et al.* (2012), and it is an extension of the random pivotal method of Deville and Tillé (1998). The basic concept is to select units in a way that the distance between them is maximised with the consequence that the sampled units are well spread in the auxiliary data space. The Euclidean distance was used in this case, but in general any distance measure can be chosen. Two different sets of auxiliary variables were applied: (1) geographical coordinates, which lead to a sample that is well spread spatially across the study area, and (2) ALS metrics, which result in a sample that is evenly spread in the space of these metrics. The idea is that the auxiliary information should be related to the target variable and thus the sample will be "a miniature version of the population" (Grafström & Schelin, 2014). An implementation of the LPM was made available via the R-package `BalancedSampling` (Grafström, 2014).

3.5.6 Sampling strategies

Three broad sampling strategies were investigated for obtaining fully integrated tree inventories: (1) single-phase sampling, (2) two-phase sampling for stratification, and (3) two-phase sampling of large clusters of population elements in the first phase and sub-sampling within selected clusters in the second phase. Inference was strictly design-based (Gregoire, 1998) in all three cases, either applying model-assisted estimation or not. The intention was to estimate total biomass for the entire population, for the three domains (F, OL, and SM), and for a union of the two TOF domains OL and SM. An overview of the sampling strategies is provided in *Figure 6*.

The population was made up of the trees in the fictitious "Skåne" population. They were selected into the sample through a 20 m × 20 m grid from which the aggregated AGB of the trees was calculated. These were assumed to be the true values. This finite grid forms the sampling frame and the single cells of the grid are the sampling units.

In total there were $N = 28{,}385{,}580$ such grid cells divided into the three domains F, SM, and OL. For each cell the domain membership, AGB, and the two ALS metrics h_{mean} and p_{veg} were known.

A standard sample size of $n = 2{,}000$ sample units was used, which corresponds approximately to the number of sub-plots used by the Swedish NFI within the county of Skåne during a five-year period. However, different to the NFI, the case of clustered field plots was not considered. The sample size was allocated proportionally to the size of the single domains resulting in $n_{\mathrm{F}} = 751$, $n_{\mathrm{TOF}} = 1{,}249$, $n_{\mathrm{OL}} = 1{,}119$, and $n_{\mathrm{SM}} = 130$ for F, TOF, OL, and SM, respectively. TOF is the union of OL and SM. The same distribution was used for all single-phase strategies that did not utilise auxiliary information, i.e. all units had the same inclusion probabilities. In cases where ALS data were incorporated for selecting samples and/or for facilitating estimation, elements with zero AGB biomass were excluded from the sampling, thus resembling stratification. Estimation was then focused on the vegetated stratum, and the sample size was reduced by the share of zero population elements to $n_{\mathrm{veg}} = 2{,}000 \times 0.516 = 968$ sample units for a fair comparison to sampling strategies that did not use auxiliary information. The main reason for this additional step was to decrease the number of elements from which to choose the sample for the LPM. Computing time for the LPM would otherwise have been too long (several days for only one repetition). The expected distribution of sample sizes was then as follows: $n_{\mathrm{F}} = 712$, $n_{\mathrm{TOF}} = 256$, $n_{\mathrm{OL}} = 181$, and $n_{\mathrm{SM}} = 75$.

Layout	Selection		Auxiliary variate	Sample size		Abbr.
single-phase	SRS		none	2 000		SRS^1
	LPM		XY	968		$\mathrm{LPM}^1_{\mathrm{XY}}$
			RS			$\mathrm{LPM}^1_{\mathrm{RS}}$
two-phase with stratification	SRS	SRS	domain	10 000	2 000	TPSS_{10}
two-phase with ALS strips	SRS	SRS	none	20	2 000	SRS^2
	LPM	SRS	XY		968	$\mathrm{LPM}_{\mathrm{XY}}$ +SRS
			RS			$\mathrm{LPM}_{\mathrm{RS}}$+SRS
		LPM	XY			$\mathrm{LPM}^2_{\mathrm{XY}}$
			RS			$\mathrm{LPM}^2_{\mathrm{RS}}$

Figure 6. Summary of the tested sampling strategies. Under **Layout**, the broad layout of the different strategies is described. **Selection** indicates which selection mechanism was applied for selecting sample units. The type of auxiliary information used during the design phase is given under **Auxiliary variate**. Here XY denotes geographical coordinates and RS stands for auxiliary information extracted from the ALS data. Under **Abbr.**, abbreviations for easier reference to the results of the simulation are given. The superscript indicates the number of sampling phases used. For LPM, the subscript shows the type of auxiliary variables in which the sample was spread. For two-phase sampling for stratification (TPSS), the subscript is related to the first-phase sample size. For the two-phase designs, the sample size for the first phase is provided in the first column under **Sample size**. The sample size for the second phase is given in the second column.

Single-phase sampling

In this case sampling units were selected directly from the entire sampling frame either through simple random sampling without replacement (SRSwoR) or the LPM. The domain total was estimated by

[3.5.2] $$\hat{t}^{(1)}_{\mathrm{HT},d} = \frac{N_d}{\hat{N}_d}\sum_{s_d}\frac{y_k}{\pi_k}.$$

The superscript (1) for $\hat{t}$ was used to indicate single-phase sampling, while the subscript d indicates a specific domain and s_d the set of sample units in the sample s that belong to the domain d. N_d is the size of domain d and $\widehat{N}_d$ its estimator. HT stands for Horvitz-Thompson estimation and was used to indicate that estimation was done without model-assistance. For both SRSwoR and the LPM, constant inclusion probabilities $\pi_k = n/N = n_{\mathrm{veg}}/N_{\mathrm{veg}}$ were used. When applying LPM sampling, either geographical coordinates or the two ALS variates h_{mean} and p_{veg} were used to select sampling units.

For model-assisted estimation, domain-specific estimation was done using

[3.5.3] $$\hat{t}^{(1)}_{\mathrm{GREG},d} = \sum_{U_d}\hat{y}_k + \frac{N_d}{\hat{N}_d}\sum_{s_d}\frac{e_k}{\pi_k}.$$

Here, the index GREG indicates that the generalised regression estimator was used for both SRSwoR and the LPM. The observed difference between observations and model predictions e_k were calculated as $e_k = y_k - \hat{y}_k$, where $\hat{y}_k$ is the model prediction from Eq. [3.5.1].

Two-phase sampling for stratification (TPSS)

In two-phase sampling for stratification, a large sample s_a of size n_a is drawn randomly and without replacement in the first phase and is stratified into $h = 1, \dots, H_{s_a}$ strata. For elements in s_a, only stratum membership is observed. The number of strata H_{s_a} is usually random and depends on s_a. Sample units with zero AGB were put into stratum $h = 1$ irrespective of land use. The remaining units were divided according to land-use domain (F, OL, or SM) into strata $h = 2, 3, \text{and } 4$, respectively. The number of first-phase units in stratum h is denoted as $n_{a,h}$. From stratum h, a sub-sample s_h of size n_h was drawn where $n_h = n_{a,h} \,/\, n_a \times n$ and $n = \sum_{h=1}^{H_{s_a}} n_h$. The latter is the sample size of the second-phase and was set to $n = 2{,}000$. The first-phase sample size was set to $n_a = 10{,}000$. The overall total was estimated as a sum of stratum totals by

[3.5.6] $$\hat{t}_{\text{TPSS}}^{(2)} = \sum_{h=1}^{H_{s_a}} \hat{t}_h^{(2)}.$$

The superscript number (2) indicates that sampling was done in two phases. For stratum totals, the following estimator was used:

[3.5.7] $$\hat{t}_{\text{HT},h}^{(2)} = N w_{a,h} \overline{y}_{s_h}.$$

Here $w_{a,h} = n_{a,h} \,/\, n_a$ is the relative size of stratum h, and $\overline{y}_{s_h}$ is the corresponding stratum mean. If model-assisted estimation was applied, the stratum total was given by

[3.5.8] $$\hat{t}_{\text{GREG},h}^{(2)} = N \frac{1}{n_a} \sum_{s_{a,h}} \hat{y}_k + N w_{a,h} \frac{1}{n_h} \sum_{s_h} e_k.$$

Two-phase sampling with ALS strips

For the second two-phase sampling design, the general inventory layout as used for a AGB estimation study in Hedmark County, Norway, was used (Ene *et al.*, 2012). The sampling frame was partitioned into $\mathrm{N_I} = 265$ non-overlapping, 500 m wide strips that ran in an east-west direction. The strips formed primary sampling units (PSUs) denoted as U_i with $i = 1, 2, \dots, N_\mathrm{I}$ such PSUs. These are clusters of population elements. Each PSU consisted of N_i secondary sampling units (SSUs) that are the actual sampling units k. In total there are $N = \sum_{U_\mathrm{I}} N_i$ SSUs and U_I denotes the entire set of PSUs. As before, sampling was done in two phases. In the first phase, a sample s_{I} of PSUs was drawn from U_{I} by some probabilistic sampling design. The sample size was chosen as $n_{\mathrm{I}} = 20$, and the selection of PSUs was performed using either SRSwoR or the LPM. In the second phase, a sample s_i of SSUs was then drawn independently from each PSU in s_{I}. Here, again, either SRSwoR or the LPM was used to select the sampling units.

In total, there were $n = \sum_{s_{\mathrm{I}}} n_i$ SSUs in the sample s, and n was set to be proportional to the size of the selected PSUs using $n_i = n_s \times (N_i / \sum_{s_{\mathrm{I}}} N_i)$.

Irrespective of the design chosen for selecting sample units in the different phases, totals were estimated as a sum of π-expanded PSU totals:

[3.5.9] $$\hat{t}_{\mathrm{C},d}^{(2)} = \sum_{s_{\mathrm{I}}} \frac{\hat{t}_{i,d}}{\pi_{\mathrm{I}}}.$$

The subscript C now indicates cluster sampling instead of strata as before, and π_{I} is the probability of including a PSU in the sample, which is constant and given as $n_{\mathrm{I}}/N_{\mathrm{I}} = 20/265$. In a similar fashion, PSU totals $\hat{t}_{i,d}$ were estimated by

[3.5.10] $$\hat{t}_{\mathrm{HT},i,d} = \sum_{s_{i,d}} y_{k|i} / \pi_{k|i}$$

where $y_{k|i}$ denotes the observed value of the study variable of unit k in the i^{th} PSU, while $s_{i,d}$ indicates summation over SSUs in the i^{th} PSU and the d^{th} domain. The probability of including an SSU in s is calculated by $\pi_{k|i} = n_i / N_i$.

In the model-assisted case, auxiliary information was needed. The PSUs were then regarded as ALS flight lines where scanning would only be done for the PSUs in s_I. The corresponding estimator for PSU totals is

[3.5.11] $$\hat{t}_{\mathrm{GREG},i,d} = \sum_{U_{i,d}} \hat{y}_k + \sum_{s_{i,d}} \frac{e_{k|i}}{\pi_{k|i}}.$$

When sampling with the LPM, again two different sets of auxiliary variables were used. For PSUs, either the geographical position of unit i in the north-south direction or the forest cover of a strip was chosen. For SSUs, the same variables as with single-phase sampling were used, i.e. the geographical coordinates or the two ALS variables h_{mean} and p_{veg}.

3.5.7 Accuracy assessment through Monte Carlo simulation

For each sampling strategy, $M = 1{,}000$ realisations were generated during the simulations. To test whether estimators were unbiased and precise, several statistics were calculated from the given sample distributions. For known population and domain totals, the generic notation t was used, and the corresponding estimates were denoted as $\hat{t}$ with a variance estimate $\widehat{V}(\hat{t})$. For the single simulation steps $m = 1, 2, \dots, M$, $\hat{t}_m$ and $\hat{V}(\hat{t}_\mathrm{m})$ were used, respectively. The variability of the total estimates was described using the observed relative standard error ($\mathrm{RSE}_\mathrm{obs}$) as calculated from the sampling distribution of $\hat{t}$:

[3.5.12] $$\mathrm{RSE}_\mathrm{obs} = \frac{\sqrt{V_\mathrm{obs}}}{t}.$$

Here, $V_\mathrm{obs} = \frac{1}{M}\sum_{m=1}^{M}\left(\hat{t}_m - \frac{1}{M}\sum_{m=1}^{M}\hat{t}_m\right)^2$. The subscript obs indicates that this is the standard error or variance as calculated from the generated sampling distribution.

4 Results

4.1 Available options for monitoring TOF

4.1.1 Definitions

The literature review revealed that currently only one general TOF definition is used at the international scale. It is the one from FAO which relies on the definitions for F and OWL (de Foresta *et al.*, 2013, p. 26).

For international reporting, the FAO definitions are accepted by most countries. The FAO defines F as land with an area of more than 0.5 ha with trees higher than 5 m and a crown cover of at least 10 %. Trees can be smaller than 5 m and the cover less than 10 % if the trees are able to reach these thresholds under the given site conditions. Lands with predominantly agricultural or urban use are not included as F or OWL regardless of the tree abundance. OWL differs from Forest with respect to the canopy cover and tree height. In the case of OWL, canopy cover should lie between 5 % and 10 % and tree height can be lower as 5 m. If shrubs are present, lands with a combined cover of trees and shrubs greater than 10 % are categorised as OWL provided that the thresholds for Forest are not reached.

By inversing the two definitions given above, TOF are trees that grow on lands with a combined cover of shrubs and trees of less than 10 % or a tree cover of less than 5 %. Patches that are smaller than 0.5 ha are always TOF independent of the crown cover. Likewise, trees on land with a predominantly agricultural or urban land use are always included as TOF.

4.1.2 NFI type inventories

Only two inventory systems were identified that include all possible types of TOF that might exist in a study area, e.g. rural and urban TOF. These are the NFMA inventories of the FAO (FAO, 2012) and the NFI of India (Tewari *et al.*, 2014).

As described earlier, the NFMA approach is a single-phase design where relatively few, but large, sampling units are distributed systematically over the study area. Trees on all land uses are measured and field plots are mapped according to changes in land use. This allows for a very straightforward implementation of the inventory and analysis of the data.

The Indian NFI, in contrast, uses a two-stage strategy (Tewari *et al.*, 2014), where the first-phase sampling units are the districts of India and a stratified random sample of 10 % is selected every 2 years proportional to the area of the physiographical zones of India. In the second stage separate sampling strategies for forest, rural TOF, and urban TOF are followed. The urban areas are divided into blocks that consist of approximately 120 to 160 households. These blocks are the sampling units, which are selected randomly. All trees in a sampled block are then measured in the field. For rural TOF, the area covered by TOF within a selected district is stratified into block, linear, and scattered tree formations using remote sensing. For each stratum a different plot layout is used which is adapted to the average spatial arrangement and density of the trees in the stratum. For the block stratum, square plots with a size of 1000 m^2 are used; for the linear stratum plots are 10 × 125 m large rectangles; and for the scattered stratum the plots are square and 3 ha large.

Beside the NFMA inventories and the Indian NFI, other countries have widened the scope of their NFIs to include TOF or use other large area monitoring strategies. Example are Switzerland (Brändli, 2010), Sweden (Fridman *et al.*, 2014), and Great Britain (Barr & Gillespie, 2000). In addition several pilot studies in the frameworks of the forest inventory and analysis programme of the USA have been conducted (Lister *et al.*, 2012; Cumming *et al.*, 2008; Riemann, 2003).

4.1.3 Remote sensing

When reviewing the literature on remote sensing in a TOF context, a distinction was made between passive (such as optical data from satellites or aerial images) and active sensors (such as light detection and ranging). The passive sensors were further separated by their spatial resolution. Data with a spatial resolution of less than 2.5 m (pixel size) was considered as high resolution. Data with a pixel size between 2.5 m and 60 m was considered as medium resolution and everything with pixels larger than 60 m as coarse spatial resolution data.

Only two studies have used data with a coarse spatial resolution to assess TOF resources (Zomer *et al.*, 2014; Perry *et al.*, 2009). Both used the vegetation continuous field product of the MODIS satellite, which delivers global maps showing the proportion of bare ground, herbaceous cover, and tree cover. The spatial resolution of the final product was 500 m. Perry *et al.* (2009) described an indirect approach for comparing land area with tree cover from MODIS with estimates from forest inventory and analysis data in the USA. The study concluded that a significant amount of tree resources is overlooked by traditional forest inventories. Zomer *et al.* (2014) combined tree cover from MODIS data with the Global Land Cover 2000 database (GLC, 2003) to assess TOF resources on agricultural land at a global scale. They found that 43 % of all agricultural land had a tree cover of more than 10 %.

At medium spatial resolution, studies typically focused either on TOF in agricultural areas or on TOF in urban settlements. The aim was often the mapping of tree cover, and images from Landsat and SPOT satellites were common data sources. A problem with this type of data is that TOF elements are often found in mixed pixels that contain spectral information about several land-use types, where standard multi-spectral classification consequently fails in recognizing them as the spectral signature of woody vegetation (Foschi & Smith, 1997). Several approaches to overcome this problem were identified: neural networks and machine vision to detect sub-pixel objects such as narrow hedgerows and single trees (Foschi & Smith, 1997); pixel swapping to assign attractiveness of sub-pixels for certain land-use classes (Thornton *et al.*, 2007; Thornton *et al.*, 2006); and spectral mixture analysis to assign fractions of certain surface types to each pixel (Small & Lu, 2006; Lee & Lathrop, 2005).

In special situations, a simple unsupervised classification approach with a subsequent supervised classification may be successful to identify woody vegetation (Kumar *et al.*, 2008). Also the prediction of carbon stocks may be an alternative when field measured carbon stocks are predicted based on specific metrics derived from the images (Myeong *et al.*, 2006).

For high spatial resolution data, a relatively clear distinction between manual interpretation of images (Fehrmann *et al.*, 2014; Walton, 2008) on the one side and an object-based classification of images on the other side (e.g., Tansey *et al.*, 2009) was found. With object-based classification, it is possible to fully map large study areas (Schumacher & Nord-Larsen, 2014; Walker & Briggs, 2007). Studies that used manual image interpretation are typically limited to some sort of sampling (Fehrmann *et al.*, 2014; Fensham & Fairfax, 2003; Hansen, 1985).

The studies on object-based classification in connection with TOF had varying objectives and target on different TOF categories. The categories can be divided into urban areas and metropolitan regions (Taubenbock *et al.*, 2010; Ouma & Tateishi, 2008; Walker & Briggs, 2007), agricultural landscapes (Liknes *et al.*, 2010; Sheeren *et al.*, 2009; Tansey *et al.*, 2009), fragmented landscapes (Zhou & Troy, 2008), suburban areas (Zhou & Troy, 2008), and savannahs (Boggs, 2010). The objectives were to map tree cover (Boggs, 2010; Liknes *et al.*, 2010; Ouma & Tateishi, 2008) but also to map other objects in a given landscape, such as buildings, streets, etc. in urban areas (Taubenbock *et al.*, 2010), and hedgerows, woodlots, crop fields, etc. in agricultural areas (Tansey *et al.*, 2009). Overall classification accuracies that were achieved ranged from 80 % to more than 90 %.

Relatively few studies on TOF monitoring apply active remote sensing techniques, such as ALS (Straub *et al.*, 2008). ALS data have been used for automated delineation and classification of forest and non-forest vegetation in urban and rural environments (Eysn *et al.*, 2012; Rutzinger *et al.*, 2008; Straub *et al.*, 2008). Straub *et al.* (2008) used full waveform ALS data to delineate regions with tree vegetation that were subsequently separated into forest and TOF. Accuracy assessments resulted in an overall accuracy of 97 %, whereas for tree groups and single trees (TOF) only 78 % and 68 % accuracy were reached, respectively. Eysn *et al.* (2012) focused on the delineation of forest areas for larger regions, which is useful from the perspective that once forest areas are identified, it implicitly becomes clear what tree cover regions can be assigned as TOF. The resulting forest maps had a high accuracy but non-forest land uses (e.g. fruit orchards) posed a challenge. Rutzinger *et al.* (2008) used ALS data to automatically detect urban trees using an object-based point cloud analysis concept. Accuracies of over 90 % were obtained, but data had a high number of returns per square meter (between 15 and 20 returns per m^2) and evaluations were made only for fairly small test sites in the city of Vienna.

Used from the ground, LiDAR data from terrestrial laser scanning appears to have a great potential to facilitate the development of volume and biomass models for TOF (Lefsky & McHale, 2008). With this technique, the entire 3D structure of trees can be revealed, and thus there is a potential to make accurate tree volume estimates from terrestrial laser scanner data alone.

4.1.4 Combining field data and remotely sensed data

In the monitoring of natural resources, there are typically two possibilities for incorporating remotely sensed data. With the first possibility, the focus is on supporting the selection of samples in the design phase of an inventory. The population may be stratified based on available maps and the sampling strategy optimised towards the specific characteristics of each stratum. This is what is done in the Indian NFI (Tewari *et al.*, 2014).

If no auxiliary data is available, the first-phase of multi-phase sampling strategy can be used to collect data, which would then be used to stratify the first-phase sample (Baffetta *et al.*, 2011; Lam *et al.*, 2011). In a second-phase, a stratified sub-sample is then selected, and one of the advantages is that sampling locations without trees can be put in a stratum that normally need not be visited in the field.

Further options supporting the selection of samples, are methods of spatially balanced sampling that strive to optimise the selection of sample units in a way that the sample represents the population as good as possible with respect to the available auxiliary information. Recently developed methods, such as the LPM, have apart from Paper IV not been applied to TOF inventories but were tested for forest inventory scenarios (Grafström *et al.*, 2014; Grafström & Ringvall, 2013).

Other approaches for combining field and remotely sensed data are directed at improving estimators in design-based inventories (Gregoire *et al.*, 2011) or at applying model-based inference (Ståhl *et al.*, 2011). In both cases variables such as aboveground biomass are modelled using predictor variables that are derived from remotely sensed data. With respect to TOF inventories, there is little to no experience about these techniques, again apart from the results of Paper IV.

4.1.5 Biomass models

For providing figures on the availability of additional wood and biomass resources, and the role of TOF in the carbon cycle, allometric models that predict biomass and carbon at the level of the individual tree are needed. For TOF, there is a general lack of such models (McHale *et al.*, 2009; Nilsson, 2008).

Only a few studies have developed TOF-specific allometric biomass equations and made a comparison to forest models (e.g., Yoon *et al.*, 2013; Kuyah *et al.*, 2012; McHale *et al.*, 2009; Zhou *et al.*, 2007; Kumar *et al.*, 1998; Nowak, 1994). In an early attempt, Nowak (1994) measured the biomass of street trees in Chicago and found that measured biomass was on average 20 % lower than the biomass predicted from forest models. In contradiction to that result, Zhou *et al.* (2014) found for open-grown trees in agricultural land, that measured biomass was on average 20 % larger than the one predicted from forest models. Adjustment factors when applying forest models to TOF may thus range between 0.8 and 1.2.

Similarly, Yoon *et al.* (2013) also partly observed an overestimation for street trees in Daegu, South Korea. However, they also found that for some species the deviation from existing forest models was small. This is in agreement with Kuyah *et al.* (2012), who found that the multi-species pan-tropical dry forest models from Chave *et al.* (2005), agreed well with models that were developed for trees in an agricultural landscape in western Kenya.

Further, Yoon *et al.* (2013) found that the application of models developed for the same species can lead to very different results. This is in agreement with the study of McHale *et al.* (2009) on street and park trees in Fort Collins, USA, which demonstrated that on single tree level the variability of biomass estimates ranged between 60 % and 300 %.

In general it can be stated that tree allometry and wood specific gravity is altered for open/grown trees in comparison to forest trees of the same species. Regarding the same species, Zhou *et al.* (2011) found, for open-grown trees, that they have a higher wood specific gravity of trunk wood, a sharper trunk taper, and that they allocate more biomass towards the crown in comparison to forest trees. They conclude that these differences are a consequence of an increased exposure of open-grown trees to solar radiation, wind, and agricultural residuals. Harja *et al.* (2012) found that open-grown trees are on average shorter than forest trees of the same species and with the same diameter.

In many countries bamboo and palms are included as TOF, and they may sometimes contribute substantially to the overall biomass stocks (FRA, 2006). For these species groups the situation regarding model availability is extremely difficult. Only some rather crude estimation approaches (e.g., Brown, 1997) were identified during the review.

4.2 Status of TOF resources in different regions of the world

In Paper II the focus was on biomass and carbon stocks; however complementary analyses focussed on other aspects such as species richness, size class distributions, and average tree size. Thus, some additional results compared to what is shown in Paper II are presented here.

4.2.1 Area estimates

As might be expected from the diversity of countries involved, cover estimates for the four major land-use classes vary to a large extent (*Table 4*). The three countries with the lowest forest cover proportion are Kyrgyzstan, Bangladesh, and Lebanon, and this implies that the majority of their land area is covered by OL, which is the land use category where TOF grow. Forest cover of more than 40 % was estimated for Zambia, Costa Rica, Honduras, and Cameroon. The other countries lie in the range of 20 % to 40 %. OWL is highest in Cameroon at 31.1 %, and most of the other countries have between 10 % and 20 % of this land-use category. Except for the countries with low forest cover, the share of OL usually ranges between 35 % and 60 %, whereas Cameron and Zambia only have about 20 % of this land-use type. Bangladesh, The Gambia, and Nicaragua have relatively high proportions of IW. Furthermore, the proportion of land where land use is unknown due to inaccessible sample units might represent up to 8 % of a country's area. In the table, also estimates of forest cover as given in the Global Forest Resources Assessment 2010 (FAO, 2010) are provided.

Table 4. *Area estimates for the major land-use classes forest, other wooded land (OWL), other land (OL), and inland water (IW) given as percentages of total land area. The corresponding standard errors are given in parentheses as a percentage of the estimated total. Under FRA 2010, the estimates of forest area as reported by the FAO in the Global Forest Resources Assessment 2010 (FAO, 2010) are given as a percentage of the total area.*

Country	Forest	FRA2010[1]	OWL	OL	IW	Unknown
Bangladesh	8.1 (18.1)	11	0.7 (62.6)	76.8 (2.9)	13.7 (12.4)	0.7 (70.6)
Cameroon	44.2 (2.6)	42	31.1 (4.2)	23.5 (3.9)	0.6 (0.3)	0.6 (0.8)
Costa Rica	46.7 (10.9)	51	1.8 (73.3)	43.1 (13.2)	4 (65.9)	4.4 (53.7)
The Gambia	26.6 (11.4)	48	10.9 (18.2)	52.1 (6.5)	10.5 (22.7)	n/a
Guatemala	37.3 (4)	34	16.3 (2.4)	42.6 (3.7)	1.8 (1.5)	2 (9.5)
Honduras	42.7 (6.9)	46	11.7 (12.1)	34.7 (7.7)	2.9 (34.7)	8 (25)
Kyrgyzstan	3.4 (14.2)	5	3.1 (14.3)	88.8 (1.1)	4.7 (15.5)	n/a
Lebanon	12.6 (12.9)	13	10.4 (13)	71.7 (3.4)	0 (76.7)	5.3 (27.9)
Nicaragua	25 (7.1)	26	17 (7.6)	48.8 (4.1)	9.2 (15.3)	n/a
Philippines	23.8 (7.9)	26	12.2 (10.1)	61.3 (3.4)	2.7 (20.6)	n/a
Zambia	63.9 (4.2)	67	7.4 (18.9)	19.7 (10.5)	4.2 (29.3)	4.8 (28.9)

[1] Estimates of forest area in the FRA 2010 report are related to total land area excluding *inland water*, whereas here estimates are related to total land area including *inland water*. The differences are partly explained by this as well as by differing inventory dates and methodologies.

4.2.2 Estimates for biomass and carbon stock

Results of living aboveground tree biomass are summarised in *Table 5*, where the average biomass per hectare, its relative standard error, and the share of the country's total biomass are presented. The largest average biomass stocks for TOF were observed in Cameroon (16.4 Mg ha^{-1}) and in the Philippines (12.3 Mg ha^{-1}). The other countries had stocks of less than 10 Mg ha^{-1} and were in general in a comparable range to the stocks found on OWL. Average biomass stocks in forests were naturally higher and ranged from 21.8 Mg ha^{-1} to 159.9 Mg ha^{-1}.

Forests typically contain the major part of the tree biomass in the countries studied. However, for six out of the eleven countries more than 10 % of the total tree biomass was found outside forests and other wooded land, and in Bangladesh as much as 75 % of the national tree biomass stocks were estimated to be TOF, mainly because OL is by far the largest land-use class in this country but also because stocks on forest land are relatively low (33.4 Mg ha^{-1}) compared to the stocks on other land (9.6 Mg ha^{-1}).

Table 5. *Estimated aboveground biomass. Given is the estimated mean aboveground biomass per hectare with the corresponding estimated relative standard error in parentheses The share of the three major land-use classes of the total tree biomass of the country is listed in the columns labelled %.*

Country	F		OWL		OL	
	Mg ha^{-1} (%)	%	Mg ha^{-1} (%)	%	Mg ha^{-1} (%)	%
Bangladesh	33.4 (21.5)	26.7	7.7 (79.5)	0.5	9.6 (8.6)	72.8
Cameroon	159.9 (2.9)	89.4	14.6 (9.9)	5.8	16.4 (15.4)	4.9
Costa Rica	104 (15.4)	93.0	0 (n/a)	0.0	8.5 (28.7)	7.0
The Gambia	21.8 (10.9)	57.6	8 (15)	8.7	6.5 (13.5)	33.7
Guatemala	80.6 (14)	86.0	9.3 (15.6)	4.3	7.9 (16.9)	9.6
Honduras	79.2 (9.3)	91.0	9.3 (16.1)	2.9	6.5 (14.5)	6.0
Kyrgyzstan	30.2 (21.8)	84.2	1 (29.8)	2.6	0.2 (22.1)	13.2
Lebanon	24.6 (28.5)	51.6	4.6 (26.5)	7.9	3.4 (26.9)	40.5
Nicaragua	74.1 (6.3)	74.4	12.6 (10.9)	8.6	8.6 (9)	17.0
Philippines	82.6 (8.2)	69.0	10.5 (12.1)	4.5	12.3 (9.2)	26.5
Zambia	32 (5.3)	95.1	4.9 (21.2)	1.7	3.6 (16.5)	3.3

4.2.3 Species richness

The general pattern for the tropical countries is that most species were found within forests followed by OL and OWL (*Table 6*). The tropical forests usually also contained the highest number of species that were only found within forests and not elsewhere, as described by the variable n_u. For Bangladesh and the two non-tropical countries of Lebanon and Kyrgyzstan, the opposite pattern was observed, i.e. more species were found in OL instead of in F, and more species were unique to OL. For other non-tropical countries, such data are not available and it would be interesting to see if this pattern can be generalised globally. Especially in Europe, it can be assumed that tree species richness is higher outside forests. However, data to evaluate this assumption are not yet available.

Table 6. *Tree species diversity by country and land use. The variables are as follows:* $\boldsymbol{S}$ *– number of species with known scientific name (only counting individuals that could be identified to the species level);* $\boldsymbol{p_{miss}}$ *– share of records with unknown species or where only family or genus were known;* $\boldsymbol{n_s}$ *– number of species with only one individual;* $\boldsymbol{n_d}$ *– number of species with two individuals;* $\boldsymbol{n_u}$ *– number of species unique within the given land use;* $\boldsymbol{p_{First}}$ *– share of the most abundant species based on tree counts;* $\boldsymbol{p_{Second}}$ *– share of the second most abundant species based on tree counts.*

Country	Land use	S	p_{miss} (%)	n_s	n_d	n_u	p_{First} (%)	Name of most abundant	p_{Second} (%)	Name of second-most abundant
Bangladesh	Forest	125	1.7	30	8	25	33.2	*Heritiera fomes*	16.2	*Excoecaria agallocha*
	OWL	32	28.4	15	8	2	15.5	*Tectona grandis*	9.7	*Olea europaea*
	OL	216	7.0	48	23	112	19.2	*Areca catechu*	9.6	*Mangifera indica*
Cameroon	Forest	435	6.1	57	30	152	3.6	*Blighia welwitschii*	2.5	*Uapaca guineensis*
	OWL	303	3.8	75	41	29	7.7	*Terminalia albida*	6.5	*Lophira lanceolata*
	OL	294	4.4	65	47	15	7.5	*Morelia senegalensis*	6.8	*Hymenocardia acida*
Costa Rica	Forest	182	32.1	61	36	118	7.2	*Tectona grandis*	3.2	*Guazuma ulmifolia*
	OL	83	23.2	42	20	19	14.3	*Cordia alliodora*	7.2	*Gliricidia sepium*
The Gambia	Forest	68	5.8	15	6	18	23.1	*Combretum glutinosum*	9.8	*Terminalia macroptera*
	OWL	42	7.2	12	3	4	23.7	*Combretum glutinosum*	17.5	*Mitragyna inermis*
	OL	58	11.3	15	8	9	9.9	*Mitragyna inermis*	9.3	*Elaeis guineensis*
Guatemala	Forest	278	29.4	83	27	98	4.7	*Pinus oocarpa*	4.1	*Pouteria reticulata*
	OWL	190	20.6	72	32	26	3.6	*Cecropia peltata*	3.2	*Byrsonima crassifolia*
	OL	234	18.2	84	45	67	5.6	*Inga spuria*	4.8	*Cordia alliodora*
Honduras	Forest	310	21.0	56	31	122	20.5	*Pinus oocarpa*	5.9	*Pinus caribaea*
	OWL	161	16.5	46	23	10	5.5	*Prosopis chilensis*	5.3	*Guazuma ulmifolia*
	OL	205	16.6	55	32	33	7.8	*Inga vera*	6.7	*Gliricidia sepium*
Kyrgyzstan	Forest	21	0.2	1	0	1	28.5	*Picea schrenkiana*	22.2	*Juniperus semiglobosa*

Country	Land use	S	p_{miss} (%)	n_s	n_d	n_u	p_{First} (%)	Name of most abundant	p_{Second} (%)	Name of second-most abundant
	OWL	18	0.7	3	2	0	49.8	*Juniperus semiglobosa*	12.9	*Acer turkestanicum*
	OL	29	0.0	3	1	8	19.1	*Populus alba*	16.6	*Sorbus tianschanica*
	Forest	33	0.7	6	3	4	21.2	*Quercus coccifera*	20.8	*Pinus brutia*
Lebanon	OWL	27	4.6	8	4	2	25.7	*Quercus coccifera*	20.4	*Juniperus excelsa*
	OL	62	20.3	10	5	29	43.1	*Olea europaea*	17.6	*Citrus*
	Forest	366	6.9	35	30	48	7.1	*Pinus caribaea*	2.7	*Guazuma ulmifolia*
Nicaragua	OWL	306	4.9	45	42	35	5.5	*Guazuma ulmifolia*	4.7	*Cecropia insignis*
	OL	338	4.1	52	33	14	8.6	*Guazuma ulmifolia*	8.0	*Cordia alliodora*
	Forest	465	43.8	139	48	190	22.6	*Shorea*	6.9	*Pentacme*
Philippines	OWL	249	22.0	82	36	32	5.2	*Gmelina arborea*	4.8	*Cocos nucifera*
	OL	285	3.6	83	35	42	57.3	*Cocos nucifera*	5.5	*Gmelina arborea*
	Forest	227	9.7	59	22	90	11.8	*Brachystegia boehmii*	7.8	*Julbernardia paniculata*
Zambia	OWL	100	15.2	34	14	5	9.5	*Diplorhynchus condylocarpon*	4.3	*Brachystegia boehmii*
	OL	120	22.2	48	21	5	6.0	*Brachystegia boehmii*	5.0	*Julbernardia paniculata*

For a considerable number of species, only one or two individuals, n_s and n_d, were observed, which is a typical feature for frequency distributions of species (Magnussen *et al.*, 2010; Lam & Kleinn, 2008). The two most abundant forest species differ in general from the two most abundant TOF species in each country examined except for Zambia. In many countries, TOF are dominated by one or two economically important species. For example, in the Philippines coconut palms (*Cocos nucifera*) account for as much as 57.3 % of all non-forest trees. Likewise in Lebanon, 60.7 % of all non-forest trees are olive trees (*Olea europaea*) and citruses (*Citrus* sp.). Similar patterns with less dominance could be found for Bangladesh, Costa Rica, The Gambia, and Kyrgyzstan. Similar observations were made by Fischer *et al.* (2011) for an NFI in Burkina Faso, where the multiple-use shea butter tree (*Vittelaria paradoxa*) clearly dominated the species frequency distribution on OL. The remaining countries showed low dominance of single species, and particularly for Cameroon, Guatemala, and Nicaragua species distributions could be considered as relatively even across all land uses.

4.2.4 Diameter distributions and average tree size

Comparing the DBH distributions for forest trees and TOF, the countries can be split into two groups. One group shows the typically inverse J-shaped distributions on both land-use categories (F and OL), with most trees occurring in the smallest size class followed by a rapid decrease in individuals with increasing diameter (*Figure 7*). The other group, i.e. The Gambia, Guatemala, the Philippines, and Zambia, shows a clear shift in the shape of the distributions from forest to OL, where the smallest size class is less frequent and medium size classes become more important. For The Gambia and the Philippines, this pattern can be attributed to the influence of palms, for which the DBH, in the absence of secondary growth, is rather constant throughout the lifespan of the tree. In Guatemala and Zambia, palms are rare and the distribution for OL might indicate a problematic regeneration of the tree population outside forests.

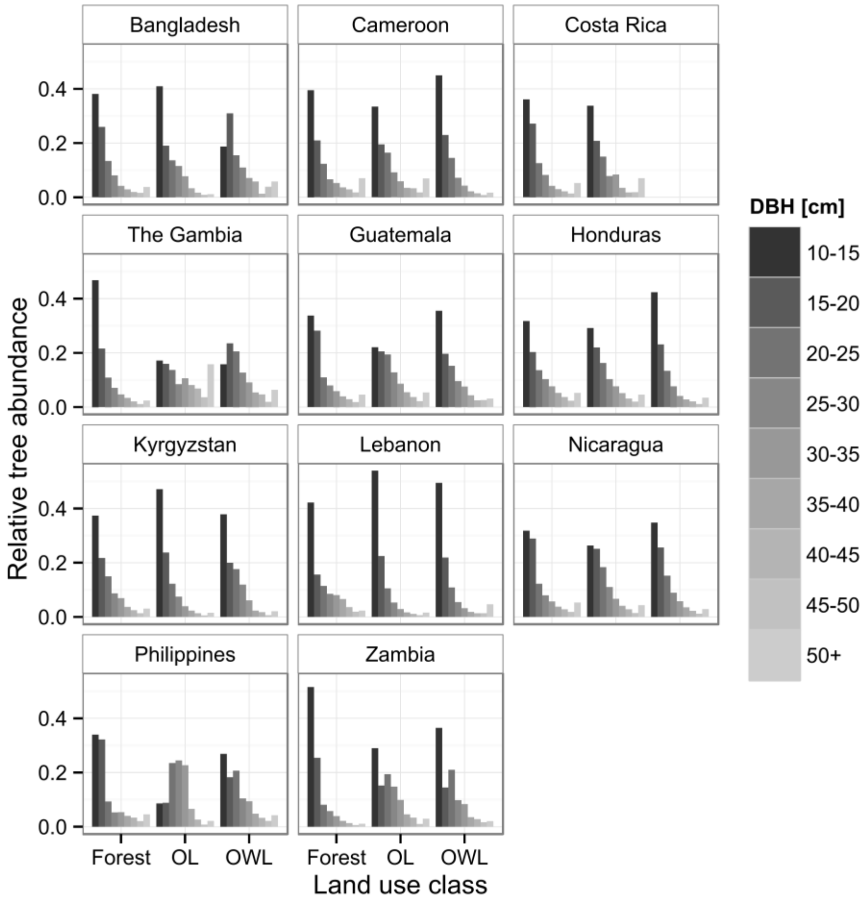

Figure 7. Relative abundance of trees based on diameter at breast height (DBH) distributions by country and major land use.

Looking at average tree size in terms of DBH and tree height, earlier observations that TOF have the tendency to be thicker in diameter and shorter in height (*Table 7*) can be confirmed. For The Gambia and the Philippines, the influence of palms is again visible because the diameters and heights outside forests are considerably larger. In Kyrgyzstan and Lebanon, TOF are thinner and shorter following a stronger shift of the species composition from forest species to fruit trees.

Table 7. *Differences in average tree size between forest and non-forest trees. Given are the diameter at breast height (DBH) in cm, total tree height in m, and the ratio between height and diameter (h/d).*

Country	Average dbh (cm)		Average height (m)		h/d ratio	
	Forest	OL	Forest	OL	Forest	OL
Bangladesh	20.7	19.7	10.4	8.7	50.06	44.29
Cameroon	22.8	23.6	13.8	9.4	60.36	39.75
Costa Rica	21.7	23.5	12.5	9.0	57.70	38.48
The Gambia	19.1	33.6	8.1	9.2	42.68	27.51
Guatemala	21.9	24.6	12.8	9.3	58.35	37.93
Honduras	23.3	23.0	13.1	8.9	56.32	38.60
Kyrgyzstan	19.8	16.9	8.8	7.0	44.41	41.64
Lebanon	21.2	17.0	7.8	4.3	36.77	25.33
Nicaragua	22.2	22.9	12.9	9.2	58.13	40.11
Philippines	21.3	26.4	10.9	12.3	51.17	46.59
Zambia	15.9	20.4	8.5	8.6	53.66	42.00

4.2.5 Distribution of TOF across land-use categories

In most cases, the majority of TOF (more than 80 %) grow on agricultural land (*Table 8*). This is supported by the area estimates where agricultural land often covers substantial parts of a country's land area. Countries that deviate from the common pattern are Bangladesh, Kyrgyzstan, and Zambia. In Bangladesh most TOF was observed in rural settlements, which in comparison to other countries also covers a large proportion of the land area. In Kyrgyzstan, the distribution of TOF across the sub-categories was relatively even, and most of the land area is covered by natural other land that consists of deserts and areas of high altitude. Other countries where a larger part of natural TOF formations such as trees on barren land, natural grassland, and marshes exist are Cameroon, The Gambia, Guatemala, and Zambia.

Table 8. *Distribution of trees outside forests across sub-categories of other land and corresponding area estimates for the sub-categories.*

Country	Trees growing on OL (percentage of total number of TOF)			Area estimates for sub-classes of OL (percentage of total land area)		
	Natural	Agriculture	Settlements	Natural	Agriculture	Settlements
Bangladesh	0	16	84	0.6	55.5	20.7
Cameroon	17.6	80.7	1.8	6.7	16	0.8
Costa Rica	0	96.4	3.6	2.7	39.2	1.2
The Gambia	17.3	82.1	0.5	8.8	39.5	3.8
Guatemala	10.9	83.2	5.9	4	37.1	1.6
Honduras	6.4	85.6	8.1	8.5	24	2.2
Kyrgyzstan	35.6	26.4	37.4	79	8	1.9
Lebanon	3.3	89.8	6.9	36.2	29.2	6.2
Nicaragua	2.7	93.5	3.8	8.5	39.2	1.1
Philippines	6.4	88.4	5.2	16.1	41.3	3.9
Zambia	30	62.7	7.3	9.5	9.5	0.7

4.3 Tree-cover patterns and their reconstruction

Here, only results for two out of the ten study sites in Paper III are shown (images 3 and 7). The original and reconstructed images are shown in *Figure 8*. Each reference image was reconstructed five times, and for all reconstructions the simulation was stopped after 1×10^6 simulation steps.

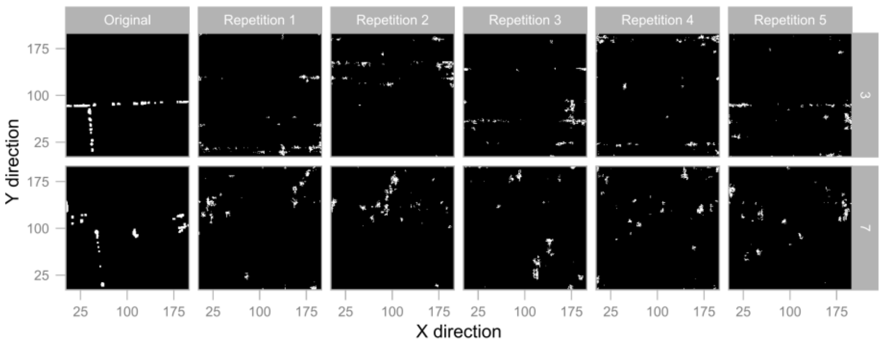

Figure 8. A selection of original and reconstructed TOF patterns. Each row represents a unique study site. The first column gives the original image extracted from the remote-sensing data. The other columns show the five replications that were simulated for each study site. The white pixels represent tree cover. Each image is 200 pixels × 200 pixels large, and one pixel has an area of 4 m².

The reconstruction algorithm (improvements-only) in combination with the tested summary statistics was in general able to reproduce images that are at least in some aspects similar to the reference images from the tested study sites. The reconstructed patterns for image 1 in the first row only partly agreed with some features from the original image. The direction of the original linear feature and the size of patches are approximately correct, but the summary functions could not identify that small patches belong to the same linear tree formation because they were not aligned to each other. This was true for all cases where linear shapes occurred.

For study sites with smaller tree groups (image 7 in *Figure 8*), the reconstruction worked well. However, for study sites with larger tree groups (not shown), results were not as good. The tendency of clustering can be recognized in the reconstructions, but the single clusters were not as clear as in the reference images. The boundaries were fuzzy and the clusters had a somewhat loose structure with many gaps.

The development of the energy during the simulations is depicted in *Figure 9*. Similar patterns were observed for all images and replications. In the beginning of the simulation, the decrease was very fast but slowed down quickly with increasing simulation time. After approximately 2×10^5 simulation steps, the energy level had reached its minimum.

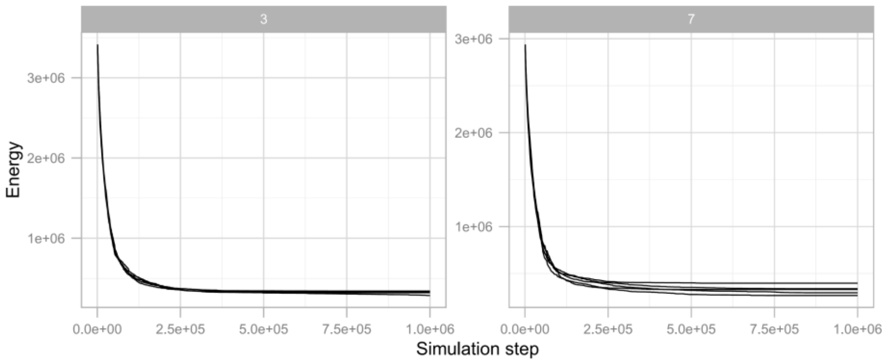

Figure 9. The development of the energy when reconstructing the two selected study sites. Each panel represents one study site, and the five lines are for the five replications.

In *Figure 10*, the lineal-path functions are given and it is shown that correlation length corresponds to the size of the largest cluster in the image.

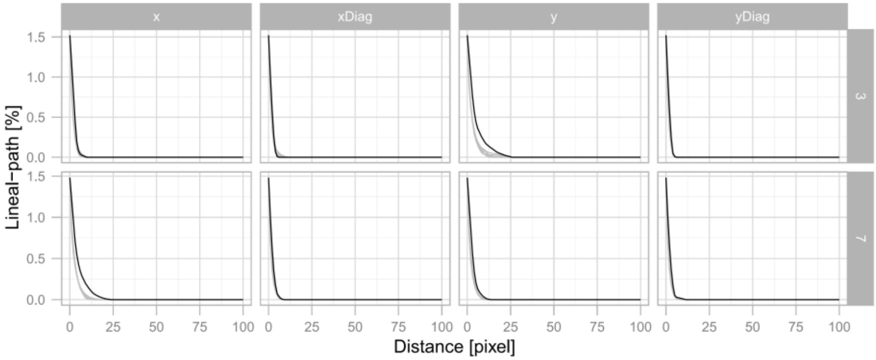

Figure 10. The lineal-path functions. Study sites are represented by the rows in the same order as in *Figure 8*. The black lines are the functions for the reference images and the grey lines are the functions for the five reconstructions of each reference image. The columns are the four directions along which functions were evaluated (from left to right: x, diagonal x, y, and diagonal y)

The lineal-path function does therefore not contain any information beyond the maximum cluster size. As a consequence, small patches that should form a linear formation did not form a line.

The two-point probability functions for the reference and reconstructed images are shown in *Figure 11*. A comparatively good agreement between reference functions and the reconstructed functions was observed. Larger deviations can occur at larger distances. Also noticeable are the deviations between the functions for the different directions; the same statistical calculation along different directions can give very different results.

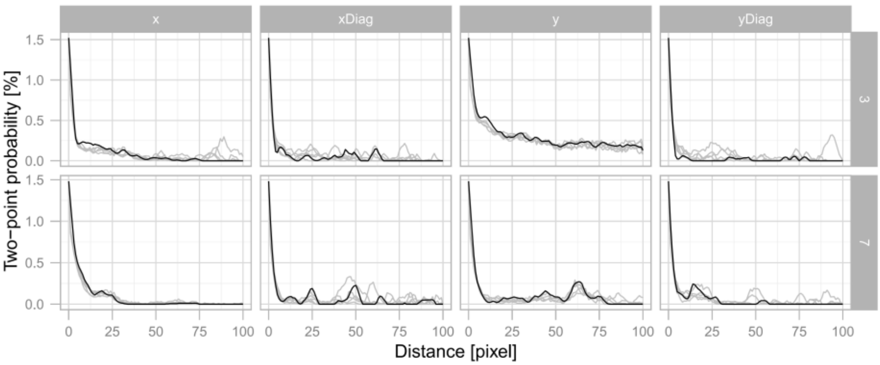

Figure 11. The two-point probability functions. Study sites are represented by the rows in the same order as in *Figure 8*. The black lines are the functions for the original images and the grey lines are the functions for the five reconstructions of each original. The columns are the four directions along which functions were evaluated (from left to right: x, diagonal x, y, and diagonal y).

4.4 Suitability of different sampling strategies

The most important findings of the simulation study are presented in the following section. Summary statistics for the artificial population that formed the basis for the simulations are given in *Table 9*. The population had a total AGB of 48.7 million Mg, including forest trees and TOF. About 12 % of that amount is found in TOF. The share of elements with no biomass at all is naturally relatively large for the TOF domain.

Table 9. *Summary for the artificial population and the different domains.*

Domain	AGB		Population elements		Zero population elements	
	Mg	%	no.	%	no.	%
Total	4.87×10^7	100.0	2.84×10^7	100.0	1.47×10^7	51.6
F	4.29×10^7	88.1	1.07×10^7	37.6	5.55×10^5	5.2
TOF	5.78×10^6	11.9	1.77×10^7	62.4	1.41×10^7	79.5
$\mathrm{TOF_{OL}}$	3.89×10^6	8.0	1.59×10^7	55.9	1.33×10^7	83.8
$\mathrm{TOF_{SM}}$	1.89×10^6	3.9	1.85×10^6	6.5	7.90×10^5	42.6

The sampling strategies aimed at estimating tree AGB for the following five domains of interest: (1) the entire tree population in Skåne (Total), (2) the forest trees (F), (3) trees outside forests (TOF), (4) trees growing in agriculture-dominated landscapes as a subset of TOF ($\mathrm{TOF_{OL}}$), and (5) trees growing in urban settings as a subset of TOF ($\mathrm{TOF_{SM}}$).

All of the presented estimators produced results that appeared to be unbiased, i.e. on average the AGB values given in *Table 9* were obtained. An exception occurred when the LPM was used to select sample units in both phases of the two-phase design with ALS strips when ALS variables were used as auxiliary information. Deviations from known totals of up to 1.94 % were observed in this case. Using ALS variables and the LPM also seemed to result in underestimation that tended to increase for smaller domains. The deviations can, however, still be considered as small.

Empirical relative standard errors are presented in *Table 10*. For single-phase sampling, the least precise results were produced when SRSwoR was applied without the use of the AGB ALS model for assisting the estimation, which was expected. The precision can nonetheless be regarded as high, especially for the entire population and the forest domain for which relative standard errors of 3.56 % and 2.56 %, respectively, of the population totals were observed. When using the AGB ALS model to assist estimation, the error was reduced considerably to 1.51 % when estimating total AGB for the entire population. A similar reduction of the sampling error was observed for the other domains as well.

For two-phase sampling for stratification, SRSwoR was used in both phases for selecting sample units. Due to the stratification of the first-phase sample, estimates were more accurate compared to single-phase SRSwoR for most of the domains or strata (*Table 10*). The precision could be improved even more by applying model-assisted estimation; in particular the smaller domains benefited from including biomass predictions into the estimators. The reduction in sampling error was, however, not as high as for single-phase sampling because auxiliary data were only available for the first-phase sampling units and not for the entire area.

When selecting ALS strips by SRSwoR, very poor precision of the applied total estimators was observed with hardly any improvement when model-assisted estimation was used (*Table 10*). Standard errors of about 15 % were observed for all domains. This was largely related to the strong variation in the size and total biomass of the ALS strips.

Table 10. *Empirical sampling errors of the simulated sampling scenarios. Presented are the standard deviations of total estimates as a percentage of true population totals (Total) or domain totals. Abbreviations under the column Design are explained in Figure 6.*

Design	Estimation design	Total	F	TOF	$\mathrm{TOF_{OL}}$	$\mathrm{TOF_{SM}}$
$\mathrm{SRS^1}$	HT	3.56	2.56	8.75	10.08	14.83
$\mathrm{SRS^1}$	GREG	1.51	1.63	4.04	4.95	7.11
$\mathrm{LPM_{XY}^1}$	HT	2.39	2.42	6.52	7.40	12.20
$\mathrm{LPM_{XY}^1}$	GREG	1.55	1.68	4.10	5.05	7.40
$\mathrm{LPM_{RS}^1}$	HT	2.38	2.27	6.63	7.68	12.59
$\mathrm{LPM_{RS}^1}$	GREG	1.58	1.69	4.02	4.96	7.19
$\mathrm{TPSS_{10}}$	HT	2.49	2.65	6.95	7.86	13.81
$\mathrm{TPSS_{10}}$	GREG	2.05	2.30	5.30	6.23	9.85
$\mathrm{SRS^2}$	HT	14.21	15.40	12.18	13.15	22.86
$\mathrm{SRS^2}$	GREG	14.45	15.62	9.81	9.72	17.35
$\mathrm{LPM_{XY} + SRS}$	HT	5.72	6.26	10.56	10.95	21.72
$\mathrm{LPM_{XY} + SRS}$	GREG	4.92	5.31	6.34	6.02	14.24
$\mathrm{LPM_{RS} + SRS}$	HT	4.37	4.91	10.42	11.32	21.95
$\mathrm{LPM_{RS} + SRS}$	GREG	3.32	3.64	7.11	6.66	15.21
$\mathrm{LPM_{XY}^2}$	HT	5.03	5.57	9.14	10.09	18.85
$\mathrm{LPM_{XY}^2}$	GREG	4.91	5.36	6.16	6.16	13.96
$\mathrm{LPM_{RS}^2}$	HT	3.38	3.83	9.57	10.44	21.26
$\mathrm{LPM_{RS}^2}$	GREG	3.22	3.58	6.93	6.56	15.34

Applying the LPM for selecting sample units generally improved the results for single-phase sampling and for two-phase sampling with ALS strips. For single-phase sampling, the reduction of standard errors was not as high as when model-assisted estimation was used in combination with SRSwoR. Applying the LPM without model assistance, the error only dropped to approximately 2.4 % independently of the choice of auxiliary variables. The use of auxiliary information in both the design and estimation stage (LPM and model-assistance) did not further reduce the sampling error compared to model-assisted SRSwoR.

In two-phase sampling with ALS strips, the design effect of using the LPM over SRSwoR was considerable and led to a three to fourfold reduction of the sampling error when estimating the population total. By using auxiliary information for selecting ALS strips, the sample of strips was well spread across the strip population and the variation in size was taken into account. For the smaller domains (OL and SM), the design effect of using the LPM was modest, while a combination with model-assisted estimation led to larger improvements. By spreading the strip sample in the ALS metrics and by applying model-assisted estimation, a sampling error that was roughly on the same level as in single-phase SRSwoR without model-assisted estimation was achieved. In particular, the smaller domains profited from incorporating the model into the estimation. Applying the LPM in both phases somewhat further reduced sampling error, but differences from the previous error were rather marginal.

5 Discussion

Since the term TOF was introduced by the FAO in 1995, considerable advancements in the large area monitoring of TOF have been made. Both FAO's NFMA inventories (FAO, 2012) and the Indian NFI (Tewari *et al.*, 2014) use a holistic approach to include all trees that grow in a country into their monitoring efforts. The NFMA inventories have a strong focus on tropical regions. For European countries, the monitoring of trees is usually centred on forests but there have been trends to widen the scope of the existing inventories to include TOF (e.g., Fridman *et al.*, 2014; Brändli, 2010). One of the drivers is to assess additional sources for sustainable wood supply. In this context, projects to initiate harmonisation of definitions and reporting on TOF have been conducted (COST, 2014). Also in the USA pilot studies for an integration of TOF into the forest inventory and analysis programme have been undertaken (e.g., Riemann, 2003). All these existing TOF inventories have shown the TOF can constitute a substantial wood resource. Further, the results of Paper II showed that the contribution to national AGB stocks may range from 3.3 % to as much as 72.8 %. Also, a strong correlation between this general importance of TOF and the forest cover of a country was identified; i.e. in countries with a relatively high forest cover, the share of TOF to national biomass stocks is low. However, in absolute terms, considerable amounts of AGB still can be accumulated outside forests. For example, the results of Paper II also showed that in Zambia only 3.3 % of tree AGB was found outside forests but these 3.3 % still accumulate approximately 52 million Mg AGB. This is more than the total tree biomass found in Kyrgyzstan or the Lebanon.

Paper II also showed that TOF are often shorter in height and thicker in diameter than forest trees, thus indicating a different tree allometry. These results are confirmed by studies from Harja *et al.* (2012) and Zhou *et al.* (2011), and lead directly to the problematic estimation of single tree biomass. As shown by the literature review (Paper I), only few studies have been conducted where TOF have been destructively sampled for the development of allometric biomass models (e.g., Zhou *et al.*, 2014 to cite the most recent one). The most common solution is to apply models that have been developed for forest trees of the same species or species group (Nair, 2012). However, results of the TOF-specific biomass studies are inconclusive and no general rule of thumb can be given regarding whether or not forest models can be transferred to TOF. Both under- and overestimation were observed, as well as cases where the forest models worked fairly well (Zhou *et al.*, 2014; Yoon *et al.*, 2013; Kuyah *et al.*, 2012; McHale *et al.*, 2009). Thus, to improve TOF models, a recommendation from the literature review is to allocate more resources to perform destructive sampling of non-forest trees to test the suitability of available forest models or to develop new models. From the point of view of this background, advances in terrestrial laser scanning for estimating the stem volume of trees (Lefsky & McHale, 2008) or other non-destructive methods for measuring volume (Rodriguez *et al.*, 2014) could be a way to avoid destructive sampling of trees and thus to reduce costs. For AGB, merely wood specific gravity would need to be determined, which ideally would be derived from small specimen sampled from the measured trees. Further, it might also be an option to add additional predictor variables to allometric models that are not as sensitive to changes in allometric scaling as height-diameter relationships are. Harja *et al.* (2012), for example, identified, the crown volume – stem diameter relationship to be relatively stable from forest to open-grown.

Another problematic area for TOF inventories is the matter of defining what TOF are, or how to distinguish trees outside forest from forest trees. As the term trees outside forests indicates, every TOF definition strongly relies on a workable forest definition. In the case of FAO's TOF definition also OWL needs to be defined as TOF are all trees outside the land-use/cover categories F and OWL. While reviewing the literature in Paper I, it was found that in particular many remote sensing studies do not clearly state which definitions they applied and how they were implemented. In contrast, the studies of Magdon *et al.* (2014) and Eysn *et al.* (2012) rigorously showed how existing land-use classification frameworks can be translated to produce land-use maps from remote sensing data. From a TOF perspective, such data are very useful since they provide a distinction of TOF from forest trees and can help in planning and conducting inventories.

A further classification of TOF into subcategories can be considered difficult because there are various criteria along which meaningful classes can be generated (Kleinn, 2000). Such classification might be done according to the land use where the trees grow, according to the geometric pattern of tree growth, according to functions (fences, wind breaks, shade, scenic beauty, etc.), or according to origin (planted or forest remnants). In a more recent FAO report, de Foresta *et al.* (2013) suggested classifying TOF into three broad sets: (1) trees growing on land with agricultural land use, (2) trees growing on land with urban land use, and (3) trees growing in small stands, in narrow linear formations, or on other lands with low crown cover. One problem with these categories is that (1) and (2) refer to land use, while (3) is based on geometrical patterns that can also be found under the categories (1) and (2).

Another fundamental problem with the way FAO's definition is interpreted is that it is seen as an area category of its own that was proposed to divide OL into mutually exclusive patches of the above-mentioned categories using relatively complicated decision trees (de Foresta *et al.*, 2013). Decision criteria like patch area and shape, tree height, canopy cover, and land use are used; however, it is often difficult to translate such criteria into operational monitoring systems. Thus, a recommendation is that TOF should not be regarded as an area category but only as a category of trees that grow on OL. If required by local conditions, local land-use classification systems can be used and the trees can be assigned to the different local categories. In some cases, trees will dominate a class (e.g., in small woodlots, hedgerows, and urban parks), while in many other cases trees will be a minor component (e.g., single trees in a field or trees along a street).

Regarding sampling strategies, the often applied approach of monitoring non-forest areas with the same sampling intensity as forest areas leads to comparatively poor precisions of TOF estimates. This was shown in Paper II and Paper IV and is a consequence of the low number of TOF elements that are selected into the sample by such single-phase sampling strategies. A large proportion of the plots on non-forest lands (OL) do not contain any trees at all. Thus it is difficult to reach high precision in the estimation of TOF attributes when the sampling designs of ordinary NFIs are expanded to include OL, except in cases where TOF are abundant and evenly distributed over the target area. However, TOF by definition occur in low densities. The precision of TOF estimates may be increased by either increasing the sample size or by increasing the plot area that is measured at each sample location.

Another option for improving the precision of sample-based estimators is to incorporate remote sensing data. If this is done using model-assisted estimation, relatively high improvements could be achieved. The study in Paper IV showed that sampling errors could be reduced by approximately 50 % and that, in particular, the TOF domains profited from including full-cover remote sensing data into the estimation. Alternatively, remote sensing data can be used to improve the selection of samples at the design stage, rather than correcting for poor samples through model-assisted estimation. In Paper IV, the LPM was applied as an option for selecting representative samples (Grafström & Schelin, 2014). Here as well, improvements of the precision were observed but not as large as when model-assisted estimation was applied. This is in contrast to Grafström and Ringvall (2013), who observed a stronger effect of the LPM in comparison to model-assisted estimation. Reasons for the different outcomes most probably can be explained by the auxiliary data used. In general, the better the auxiliary data explains the target variable, the better the estimation will be. With respect to tree inventories, the LPM has, apart from Paper IV, so far only been tested for forest inventories (Grafström *et al.*, 2014; Grafström & Ringvall, 2013). More simulation studies are thus needed to support any general conclusions about this relatively novel sampling design.

Instead of single-phase sampling strategies, multi-phase sampling strategies may be applied. In Paper IV, two-phase sampling for stratification estimators showed a higher precision than their single-phase counterparts when model-assisted sampling was not used. The model-assisted estimators were, however, slightly less precise because auxiliary information was only available for the first-phase sample. Multi-phase sampling strategies have also been applied in other studies of TOF (Baffetta *et al.*, 2011; Corona *et al.*, 2011; Lam *et al.*, 2011; Baffetta *et al.*, 2009). They have the advantage that information can be gathered for a large sample during the first-phase, which is especially useful when little or nothing is known about the objects of interest. This is often the case for TOF because an extraction and classification from remote sensing data can be difficult and requires data of high spatial resolution.

For ALS assisted forest inventories, a common approach is to use a two-phase strategy, where a sample of ALS strips is selected in the first phase to reduce the acquisition costs of the ALS data (Gobakken *et al.*, 2012). Results of Paper IV indicated that this strategy is less optimal for TOF inventories because of the clustering of sample units within the ALS strips. In single-phase strategies and with two-phase sampling for stratification, sampling units are more evenly spread over the entire study area, thus resulting in lower sampling errors. In particular, when ALS strips were selected randomly, poor precisions of estimators were obtained due to considerable variability in strip size.

Regarding Paper III, it has already been mentioned that the reconstruction technique could not be used as intended for simulating maps of TOF cover. Further, the reconstructed patterns of canopy cover did only partly resemble the original patterns as extracted from the ALS data. Future studies could focus on four possible approaches for improving the reconstruction: (1) the use of other initial configurations, e.g. cluster processes; (2) alternative ways for randomly selecting pixels for random movement; (3) additional or different summary statistics; and (4) including Metropolis-Hastings probabilities (Metropolis *et al.*, 1953).

The initial configuration for each simulation was a purely random image. In order to improve the reconstructed images, start configurations could be initialised in a way that random linear structures or random clusters are produced. To do this, techniques from point pattern statistics could be applied, including the Neyman-Scott process for clustering (Illian *et al.*, 2008, chap. 6.3) or fibre processes for linear features (Stoyan *et al.*, 1995, chap. 9).

To prevent the rapid decrease in energy, Jiao et al. (2008) suggested only selecting *free* pixels, i.e. pixels that are not surrounded by others of the same type, at later stages of the simulation. Jiao *et al.* (2008) showed that this approach helps to further decrease the energy and increase the similarity between the reconstructed images and the original images with respect to the summary statistics used in the energy function. It is also likely that this technique helps to make single patches of trees more compact.

Further, pixel selection might be based on landscape features such as roads or watercourses along which TOF often occur. In Paper III, pixels were selected with equal probabilities, but as an alternative the probabilities could be modified in such a way that locations close to roads and watercourses are preferred. The reconstructed images would still be random representations of the original images.

The third approach for improvements involves alternative summary statistics that capture more information from the reference patterns. Jiao *et al.* (2009) described a superior descriptor for random textures that uses a two-point cluster function to evaluate whether two points separated by a distance are situated in the same cluster. This summary statistic is superior in terms of describing topological connectivity of single patches and leads to more accurate results. In a similar manner, the lineal-path function could be modified in a way that it checks whether two points in different clusters are aligned to each other. Other summary statistics that can be used for reconstruction purposes are described in Schlüter and Vogel (2011). However, many of the alternatives are computationally intensive and cannot be easily applied.

Another reason for the fast decrease of the energy could be that the simulations got stuck in local minima because we used an *improvements-only* algorithm. The Metropolis-Hastings probability was originally suggested by Kirkpatrick *et al.* (1983) because it allows an occasional acceptance of inferior energy values. That is, a new configuration can be accepted even if the resulting energy is larger. Tscheschel and Stoyan (2006) showed that the *improvements-only* algorithm was able to produce results with small deviations from the original patterns in a shorter amount of time. The difficulties with using the Metropolis-Hastings probabilities lie in the determination of the parameters that are needed to calculate them. In this study, initial tests with this approach showed no improvements.

6 Conclusions and future work

The literature review showed that methodologies for large area tree inventories are available. For designing and implementing such inventories, the rich experience from the forest inventory sector can be used, and it is thus a natural choice to assess TOF as part of existing forest inventory systems. This is increasingly being done, but only in few cases the specific features of TOF variables have been carefully considered in the sampling strategies. One important example is the Indian NFI. In addition, only rarely have all possible TOF types been included in existing inventories, and only the Indian NFI and the NFMA inventories of the FAO include all types of trees that might grow in a country. In general, when it comes to the reporting and the analysis of the results, harmonisation is not as advanced as it is for the forestry sector. Also, it is often difficult to locate the results from TOF assessments. Coordination in this area is poor, although the FAO has taken several steps towards mainstreaming TOF into inventories and the resulting databases.

Technical aspects that need to be solved include allometric biomass models, which are a bottleneck for all studies related to biomass and carbon stocks. Huge uncertainties exist, in particular when forest models are applied in the absence of TOF-specific models. More efforts for developing models and for assessing the adequacy of existing forest models for TOF assessments are urgently needed. In particular, new methods for non-destructive estimations of tree volume using terrestrial laser scanning are promising because they have the potential to reduce laborious fieldwork.

For effective TOF inventories, an extensive use of remotely sensed data to assist field inventories appears to be necessary. However, further research to investigate schemes such as model-assisted estimation or recently developed methods for spatially balanced sampling is needed.

The analysis of NFMA data provided quantitative evidence that TOF are an important tree resource on the national level. In many countries, TOF are a substantial part of the national woody biomass stocks, and TOF might even be an important regional wood resource.

Reconstruction techniques that were originally developed for applications in material physics appear to have promise for verifying the applicability of summary statistics for describing tree cover in the open landscape. In the long run, this technique might also help to generate artificial populations of tree cover with different spatial properties. Such populations can be used in sampling simulations for identifying appropriate inventory strategies.

The sampling simulation study showed that several promising strategies for integrating TOF into NFIs exist. In particular, when model-assisted estimation was used, the most straightforward way of extending a basic sampling design led to good results. This, however, requires the availability of wall-to-wall remotely sensed data. To circumvent this requirement, two-phase sampling strategies can be applied. Although two-phase sampling for stratification led to slightly larger sampling errors when compared to the single-phase model-assisted strategies, this seems to be a fair trade-off for cases where wall-to-wall auxiliary data are not available. In addition, detailed information about domain membership can be collected for populations that are difficult to classify beforehand, such as TOF. When using two-phase sampling with ALS strips, extensive use of auxiliary information at the design and estimation stage was required to reach similar precision as in single-phase random selection of sampling units without any use of auxiliary information. A final conclusion is that the use of remotely sensed auxiliary data appears to be very important for setting up cost-efficient sampling strategies and that sample errors can be reduced substantially by correctly incorporating such data both at the estimation stage and at the design stage of TOF surveys.

References

Ahmed, P. (2008). Trees outside forests (TOF): a case study of wood production and consumption in Haryana. *International Forestry Review,* 10, pp. 165-172.

Alexandre, D.Y., Lescure, J.P., Bied-Charreton, M. & Fotsing, J.M. (1999). *Contribution à l'état des connaissances sur les arbres hors forêt (TOF)*. Orléans: IRD-FAO.

Baffetta, F., Corona, P. & Fattorini, L. (2011). Assessing the attributes of scattered trees outside the forest by a multi-phase sampling strategy. *Forestry,* 84, pp. 315-325.

Baffetta, F., Fattorini, L. & Corona, P. (2009). Estimation of small woodlot and tree row attributes in large-scale forest inventories. *Environmental and Ecological Statistics,* 18, pp. 147-167.

Barr, C.J. & Gillespie, M.K. (2000). Estimating hedgerow length and pattern characteristics in Great Britain using Countryside Survey data. *Journal of Environmental Management,* 60, pp. 23-32.

Baudry, J., Bunce, R.G.H. & Burel, F. (2000). Hedgerows: An international perspective on their origin, function and management. *Journal of Environmental Management,* 60, pp. 7-22.

Bäuerle, H. & Nothdurft, A. (2011). Spatial modeling of habitat trees based on line transect sampling and point pattern reconstruction. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere,* 41(4), pp. 715-727.

Bellefontaine, R., Petit, S., Deleporte, P. & Bertault, J.-G. (2002). Trees outside forests. Towards better awareness., p. 218.

Bélouard, T. & Coulon, F. (2002). Trees outside forests: France. In: Bellefontaine, R., Petit, S., Deleporte, P. & Bertault, J.-G. (eds) *Trees outside forests. Towards better awareness.* Rome: FAO, pp. 149-156.

Bhagwat, S.A., Willis, K.J., Birks, H.J.B. & Whittaker, R.J. (2008). Agroforestry: a refuge for tropical biodiversity? *Trends in Ecology & Evolution,* 23, pp. 261-267.

Bieberstein, D., Pfaff, K., Wünsche, H.E. & Grossmann, H. (1975). Flurholzinventur in ausgewählten Oberförstereien der Deutschen Demokratischen Republik (Inventory of timber growing outside the forest in selected forestry districts in E. Germany. *Sozialistische Forstwirtschaft*, 25, pp. 48-49.

Bieberstein, D., Wernicke, F., Krausewald, G., Joachim, H.F. & Tesch, U. (1982). Inventur und Planung der Bewirtschaftung der Flurgehölze - eine wichtige Grundlage für die Erschließung von Reserven (Inventory and management planning of timber growing outside the forest - an important basis making these reserves available). *Sozialistische Forstwirtschaft,* 32, pp. 353-355.

Boffa, J.M. (2000). West African agroforestry parklands: keys to conservation and sustainable management. *Unasylva 200,* 51, pp. 11-17.

Boggs, G.S. (2010). Assessment of SPOT 5 and QuickBird remotely sensed imagery for mapping tree cover in savannas. *International Journal of Applied Earth Observation and Geoinformation,* 12, pp. 217-224.

Bowler, D.E., Buyung-Ali, L., Knight, T.M. & Pullin, A.S. (2010). Urban greening to cool towns and cities: A systematic review of the empirical evidence. *Landscape and Urban Planning,* 97, pp. 147-155.

Brändli, U.-B. (2010). *Schweizerisches Landesforstinventar. Ergebnisse der dritten Erhebung 2004-2006*. Birmensdorf, Bern: WSL, BAFU.

Brown, S. (1997). *Estimating biomass and biomass change of tropical forests: a primer*. (FAO Forestry Paper, 137). Rome: FAO.

Carlowitz, H.C.v. (1713). *Sylvicultura oeconomica: Anweisung zur wilden Baum-Zucht.* Leipzig: Braun.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Folster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.P., Nelson, B.W., Ogawa, H., Puig, H., Riera, B. & Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia,* 145(1), pp. 87-99.

Chave, J., Rejou-Mechain, M., Burquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martinez-Yrizar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Pelissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G. & Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology,* 20(10), pp. 3177-3190.

Corona, P., Agrimi, M., Baffetta, F., Barbati, A., Chiriacò, M.V., Fattorini, L., Pompei, E., Valentini, R. & Mattioli, W. (2011). Extending large-scale forest inventories to assess urban forests. *Environmental monitoring and assessment*.

COST *Action FP1001: Improving Data and Information on the Potential Supply of Wood Resources: A European Approach from Multisource National Forest Inventories (USEWOOD).* http://www.cost.eu/COST_Actions/fps/Actions/FP1001 [15/10/2014].

Cumming, A.B., Nowak, D.J., Twardus, D., Hoehn, R., Mielke, M. & Rideout, R. (2007). *National Forest Health Monitoring Program, Urban Forests Of Wisconsin: Pilot Monitoring Project 2002*. Newton Square, PA: USDA, Forest Service, .

Cumming, A.B., Twardus, D.B. & Nowak, D.J. (2008). Urban forest health monitoring: Large scale assessments in the United States. *Arboriculture and Urban Forestry,* 34, pp. 341-346.

de Foresta, H., Somarriba, E., Temu, A., Boulanger, D., Feuilly, H. & Gaulthier, M. (2013). *Towards the Assessment of Trees Outside of Forests*. Rome.

Deville, J.-C. & Tillé, Y. (1998). Unequal Probability Sampling Without Replacement Through a Splitting Method. *Biometrika,* 85, pp. 89-101.

Eddelbuettel, D. (2013). *Seamless R and C++ Integration with Rcpp*. New York: Springer.

Ene, L.T., Næsset, E., Gobakken, T., Gregoire, T.G., Ståhl, G. & Nelson, R. (2012). Assessing the accuracy of regional LiDAR-based biomass estimation using a simulation approach. *Remote Sensing of Environment,* 123, pp. 579-592.

Eysn, L., Hollaus, M., Schadauer, K. & Pfeifer, N. (2012). Forest Delineation Based on Airborne LIDAR Data. *Remote Sensing,* 4, pp. 762-783.

FAO (2010). *Global Forest Resource Assessment 2010. Main Report*. Rome: FAO.

FAO (2012). *National Forest Monitoring and Assessment – Manual for integrated field data collection. Version 3.0*. (NFMA Working Paper, 37/E). Rome, Italy: FAO.

Fehrmann, L., Seidel, D., Krause, B. & Kleinn, C. (2014). Sampling for landscape elements-a case study from Lower Saxony, Germany. *Environmental monitoring and assessment,* 186(3), pp. 1421-1430.

Fensham, R.J. & Fairfax, R.J. (2003). Assessing woody vegetation cover change in north-west Australian savanna using aerial photography. *International Journal of Wildland Fire,* 12, pp. 359-367.

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, pp. 1532-1540.

Foschi, P.G. & Smith, D.K. (1997). Detecting subpixel woody vegetation in digital imagery using two artificial intelligence approaches. *Photogrammetric engineering and remote sensing,* 63, pp. 493-499.

Frangi, J.L. & Lugo, A.E. (1985). Ecosystem dynamics of a subtropical floodplain forest. *Ecological Monographs,* 55(3), pp. 351-369.

Fridman, J., Holm, S., Nilsson, M., Nilsson, P., Ringvall, A. & Ståhl, G. (2014). Adapting National Forest Inventories to changing requirements – the case of the Swedish National Forest Inventory at the turn of the 20th century. *Silva Fennica,* 48, pp. 1-29.

FSI (2011). *India State of Forest Report*. Dehradun, India: Forest Survey of India.

Gerke, K.M., Karsanina, M.V., Vasilyev, R.V. & Mallants, D. (2014). Improving pattern reconstruction using directional correlation functions. *Epl,* 106(6).

GLC The Global Land Cover Map for the Year 2000, 2003. GLC 2000 database. In: 2003.

Gobakken, T., Næsset, E., Nelson, R., Bollandsås, O.M., Gregoire, T.G., Ståhl, G., Holm, S., Ørka, H.O. & Astrup, R. (2012). Estimating biomass in Hedmark County, Norway using national forest inventory field plots and airborne laser scanning. *Remote Sensing of Environment,* 123, pp. 443-456.

Grafström, A. (2014). *BalancedSampling: Balanced and spatially balanced sampling*. (Version: R package version 1.4) [Computer Program]. Available from: http://cran.r-project.org/package=BalancedSampling.

Grafström, A., Lundström, N.L.P. & Schelin, L. (2012). Spatially Balanced Sampling through the Pivotal Method. *Biometrics,* 68(2), pp. 514-520.

Grafström, A. & Ringvall, A.H. (2013). Improving forest field inventories by using remote sensing data in novel sampling designs. *Canadian Journal of Forest Research,* 43, pp. 1015-1022.

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

Grafström, A. & Schelin, L. (2014). How to Select Representative Samples. *Scandinavian Journal of Statistics,* 41, pp. 277-290.

Grala, R.K., Tyndall, J.C. & Mize, C.W. (2010). Impact of field windbreaks on visual appearance of agricultural lands. *Agroforestry Systems,* 80, pp. 411-422.

Gregoire, T.G. (1998). Design-based and model-based inference in survey sampling: appreciating the difference. *Canadian Journal of Forest Research,* 28(10), pp. 1429-1447.

Gregoire, T.G., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R. & Holm, S. (2011). Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway. *Canadian Journal of Forest Research,* 41, pp. 83-95.

Hansen, M.H. (1985). Notes: Line Intersect Sampling of Wooded Strips. *Forest Science,* 31, pp. 282-288.

Harja, D., Vincent, G., Mulia, R. & van Noordwijk, M. (2012). Tree shape plasticity in relation to crown exposure. *Trees,* 26, pp. 1275-1285.

Herzog, F. (1998). Streuobst: a traditional agroforestry system as a model for agroforestry development in temperate Europe. *Agroforestry Systems,* 42(1), pp. 61-80.

Herzog, F. (2000). The importance of perennial trees for the balance of northern European agricultural landscapes. *Unasylva 200,* 51, pp. 42-48.

Hinsley, S.A. & Bellamy, P.E. (2000). The influence of hedge structure, management and landscape context on the value of hedgerows to birds: A review. *Journal of Environmental Management,* 60(1), pp. 33-49.

Höfle, B., Mücke, W., Dutter, M., Rutzinger, M. & Dorninger, P. (2009). Detection of building regions using airborne LiDAR - A new combination of raster and point cloud based GIS methods. In: Car, A., Griesebner, G. & Strobl, J. (eds) *Crossroads @ GI_Forum' 09: Proceedings of the Geoinformatics Forum Salzburg.* Salzburg, pp. 66-75.

Holmgren, P., Masakha, E.J. & Sjöholm, H. (1994). Not all African land is being degraded: a recent survey of trees on farms in Kenya reveals rapidly increasing forest resources. *Ambio,* 23, pp. 390-395.

Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. Chichester: John Wiley & Sons.

IPCC (2000). *Land Use, Land-Use Change, and Forestry: A Special Report of the Intergovernmental Panel on Climate Change*. Cambridge: IPCC.

IPCC (2006). *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme*. Japan: IGES.

Jiao, Y., Stillinger, F.H. & Torquato, S. (2008). Modeling heterogeneous materials via two-point correlation functions. II. Algorithmic details and applications. *Physical Review E,* 77, p. 15.

Jiao, Y., Stillinger, F.H. & Torquato, S. (2009). A superior descriptor of random textures and its predictive capacity. *Proceedings of the National Academy of Sciences,* 106, pp. 17634-17639.

Jim, C.Y. & Chen, W.Y. (2009). Ecosystem services and valuation of urban forests in China. *Cities,* 26(4), pp. 187-194.

Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P. (1983). Optimization by simulated annealing. *Science,* 220, pp. 671-680.

Kleinn, C. (1999). *Compilation of Information on Trees Outside the Forest. A contribution to the Forest Resources Assessment 2000 of FAO. Regional Special Study for Latin America.* Costa Rica: CATIE.

Kleinn, C. (2000). On large-area inventory and assessment of trees outside forests. *Unasylva 200,* 51, pp. 3-10.

Konijnendijk, C.C. (2003). A decade of urban forestry in Europe. *Forest Policy and Economics,* 5, pp. 173-186.

Kumar, A., Singh, K., Lal, B. & Singh, R. (2008). Mapping of apple orchards using remote sensing techniques in cold desert of Himachal Pradesh, India. *Journal of the Indian Society of Remote Sensing,* 36, pp. 387-392.

Kumar, B.M., Suman, J.G., Jamaludheen, V. & Suresh, T.K. (1998). Comparison of biomass production, tree allometry and nutrient use efficiency of multipurpose trees grown in woodlot and silvopastoral experiments in Kerala, India. *Forest Ecology and Management,* 112, pp. 145-163.

Kuyah, S., Dietz, J., Muthuri, C., Jamnadass, R., Mwangi, P., Coe, R. & Neufeldt, H. (2012). Allometric equations for estimating biomass in agricultural landscapes: I. Aboveground biomass. *Agriculture, Ecosystems & Environment,* 158, pp. 216-224.

Lam, T.Y. & Kleinn, C. (2008). Estimation of tree species richness from large area forest inventory data: Evaluation and comparison of jackknife estimators. *Forest Ecology and Management,* 255, pp. 1002-1010.

Lam, T.Y., Kleinn, C. & Coenradie, B. (2011). Double sampling for stratification for the monitoring of sparse tree populations: the example of Populus euphratica Oliv. forests at the lower reaches of Tarim River, Southern Xinjiang, China. *Environmental monitoring and assessment,* 175, pp. 45-61.

Lantmäteriet *Fastighetskartan*. http://www.lantmateriet.se/sv/Kartor-och-geografisk-information/Kartor/Fastighetskartan/ [2014-12-02].

Lantmäteriet *GSD-Höjddata, grid 2+*. http://www.lantmateriet.se/sv/Kartor-och-geografisk-information/Hojddata/GSD-Hojddata-grid-2/ [2014-12-02].

Lantmäteriet *Översiktskartan*. http://www.lantmateriet.se/sv/Kartor-och-geografisk-information/Kartor/Oversiktskartan/ [2014-12-02].

Lee, S. & Lathrop, R.G. (2005). Sub-pixel estimation of urban land cover components with linear mixture model analysis and Landsat Thematic Mapper imagery. *International Journal of Remote Sensing,* 26, pp. 4885-4905.

Lefsky, M. & McHale, M.R. (2008). Volume estimates of trees with complex architecture from terrestrial laser scanning. *Journal of Applied Remote Sensing,* 2, p. 023521.

Liknes, G.C., Perry, C.H. & Meneguzzo, D.M. (2010). Assessing tree cover in agricultural landscapes using high-resolution aerial imagery. *Journal of Terrestrial Observation,* 2, pp. 38-55.

Lilleleht, a., Sims, A. & Pommerening, A. (2013). Spatial forest structure reconstruction as a strategy for mitigating edge-bias in circular monitoring plots. *Forest Ecology and Management*.

Lister, A.J., Scott, C.T. & Rasmussen, S. (2012). Inventory methods for trees in nonforest areas in the Great Plains States. *Environmental monitoring and assessment,* 184, pp. 2465-74.

Lund, H.G. (2002). When Is a Forest Not a Forest? *Journal of Forestry,* 100, pp. 21-28.

Magdon, P., Fischer, C., Fuchs, H. & Kleinn, C. (2014). Translating criteria of international forest definitions into remote sensing image analysis. *Remote Sensing of Environment,* 149, pp. 252-262.

Magnussen, S., Smith, B., Kleinn, C. & Sun, I.F. (2010). An urn model for species richness estimation in quadrat sampling from fixed-area populations. *Forestry,* 83, pp. 293-306.

Manning, A.D., Gibbons, P. & Lindenmayer, D.B. (2009). Scattered trees: a complementary strategy for facilitating adaptive responses to climate change in modified landscapes? *Journal of Applied Ecology,* 46, pp. 915-919.

Marklund, L.G. (1988). *Biomass functions for pine, spruce and birch in Sweden.* Umeå: Dep. of For. Surv.

McGarigal, K., Cushman, S.A., Neel, M.C. & Ene, E. (2012). *FRAGSTATS v4: Spatial Pattern Analysis Program for Categorial and Continuous Maps.* Amherst: University of Massachusetts. Available from: http://www.umass.edu/landeco/research/fragstats/fragstats.html.

McHale, M.R., Burke, I.C., Lefsky, M.A., Peper, P.J. & McPherson, E.G. (2009). Urban forest biomass estimates: is it important to use allometric relationships developed specifically for urban trees? *Urban Ecosystems,* 12, pp. 95-113.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics,* 21, p. 1087.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. & New, M. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100). In: 2004, p. 25.

Myeong, S., Nowak, D.J. & Duggin, M.J. (2006). A temporal analysis of urban forest carbon storage using remote sensing. *Remote Sensing of Environment,* 101, pp. 277-282.

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment,* 80, pp. 88-99.

Nair, P.K.R. (2012). Carbon sequestration studies in agroforestry systems: a reality-check. *Agroforestry Systems,* 86(2), pp. 243-253.

Nilsson, S. (2008). The Indian forestry system at a crossroads: an outsider’s view. *International Forestry Review,* 10, pp. 414-421.

Nothdurft, A., Saborowski, J., Nuske, R.S. & Stoyan, D. (2010). Density estimation based on k-tree sampling and point pattern reconstruction. *Canadian Journal of Forest Research,* 40(5), pp. 953-967.

Nowak, D. (2002). Carbon storage and sequestration by urban trees in the USA. *Environmental Pollution,* 116, pp. 381-389.

Nowak, D.J. (1994). Atmospheric Carbon Dioxide Reduction by Chicago's Urban Forest. In: McPherson, E.G., Nowak, D.J. & Rowntree, R.A. (eds) *Chicago's Urban Forest Ecosystem: Results of the Chicago Urban Forest Climate Project.* Radnor, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, pp. 83-94.

Nowak, D.J., Crane, D.E., Stevens, J.C., Hoehn, R.E., Walton, J.T. & Bond, J. (2008). A ground-based method of assessing urban forest structure and ecosystem services. *Arboriculture and Urban Forestry,* 34, pp. 347-358.

Ouma, Y.O. & Tateishi, R. (2008). Urban-trees extraction from Quickbird imagery using multiscale spectex-filtering and non-parametric classification. *ISPRS Journal of Photogrammetry and Remote Sensing,* 63(3), pp. 333-351.

Pain-Orcet, M. & Bellefontaine, R. (2004). Trees outside the forest: A new perspective on the management of forest resources in the tropics. In: Babin, D. (ed. *Beyond Tropical Deforestation*. Paris: UNESCO/CIRAD, pp. 423-430.

Paletto, A. & Chincarini, M. (2012). Heterogeneity of linear forest formations: differing potential for biodiversity conservation. A case study in Italy. *Agroforestry Systems,* 86, pp. 83-93.

Pandey, D. (2008). Trees outside the forest (TOF) resources in India. *International Forestry Review,* 10(2), pp. 125-133.

Perry, C.H., Woodall, C.W., Liknes, G.C. & Schoeneberger, M.M. (2009). Filling the gap: improving estimates of working tree resources in agricultural landscapes. *Agroforestry Systems,* 75, pp. 91-101.

Plieninger, T. (2011). Capitalizing on the Carbon Sequestration Potential of Agroforestry in Germany's Agricultural Landscapes: Realigning the Climate Change Mitigation and Landscape Conservation Agendas. *Landscape Research,* 36, pp. 435-454.

Pommerening, A. (2006). Evaluating structural indices by reversing forest structural analysis. *Forest Ecology and Management,* 224(3), pp. 266-277.

Pommerening, A. & Stoyan, D. (2008). Reconstructing spatial tree point patterns from nearest neighbour summary statistics measured in small subwindows. *Canadian Journal of Forest Research,* 38(5), pp. 1110-1122.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. [Computer Program]. Vienna, Austria: R Foundation for Statistical Computing. Available from: http://www.R-project.org/.

Riemann, R. (2003). *Pilot Inventory of FIA plots Traditionally Called "Nonforest"*. Newton Square, PA: USDA, Forest Service.

Rodriguez, F., Lizarralde, I., Fernandez-Landa, A. & Condes, S. (2014). Non-destructive measurement techniques for taper equation development: a study case in the Spanish Northern Iberian Range. *European Journal of Forest Research,* 133(2), pp. 213-223.

Rutzinger, M., Höfle, B., Hollaus, M. & Pfeifer, N. (2008). Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification. *Sensors,* 8, pp. 4505-4528.

Rydberg, D. & Falck, J. (2000). Urban forestry in Sweden from a silvicultural perspective: A review. *Landscape and Urban Planning,* 47(1-2), pp. 1-18.

Saldarriaga, J.G., West, D.C., Tharp, M.L. & Uhl, C. (1988). Long-Term Chronosequence of Forest Succession in the Upper Rio Negro of Colombia and Venezuela. *Journal of Ecology,* 76(4), pp. 938-958.

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

SCBD (2005). *Handbook of the Convention on Biological Diversity Including its Cartagena Protocol on Biosafety*. 3rd. ed. Montreal, Canada: CBD, UNEP.

Schlüter, S. & Vogel, H.-J. (2011). On the reconstruction of structural and functional properties in random heterogeneous media. *Advances in Water Resources,* 34, pp. 314-325.

Schnell, S., Altrell, D., Ståhl, G. & Kleinn, C. (2015). The contribution of trees outside forests to national tree biomass and carbon stocks-a comparative study across three continents. *Environmental monitoring and assessment,* 187, p. 4197.

Schoeneberger, M., Bentrup, G., de Gooijer, H., Soolaneyakanahally, R., Sauer, T., Brendle, J., Zhou, X.H. & Current, D. (2012). Branching out: Agroforestry as a climate change mitigation and adaptation tool for agriculture. *Journal of Soil and Water Conservation,* 67(5), pp. 128a-136a.

Schoeneberger, M.M. (2009). Agroforestry: working trees for sequestering carbon on agricultural lands. *Agroforestry Systems,* 75(1), pp. 27-37.

Schumacher, J. & Nord-Larsen, T. (2014). Wall-to-wall tree type classification using airborne lidar data and CIR images. *International Journal of Remote Sensing,* 35(9), pp. 3057-3073.

Sheeren, D., Bastin, N., Ouin, A., Ladet, S., Balent, G. & Lacombe, J.-P. (2009). Discriminating small wooded elements in rural landscape from aerial photography: a hybrid pixel/object-based analysis approach. *International Journal of Remote Sensing,* 30, pp. 4979-4990.

Small, C. & Lu, J.W.T. (2006). Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis. *Remote Sensing of Environment,* 100(4), pp. 441-456.

Ståhl, G., Holm, S., Gregoire, T.G., Gobakken, T., Næsset, E. & Nelson, R. (2011). Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. *Canadian Journal of Forest Research,* 41, pp. 96-107.

Stoyan, D., Kendall, W.S. & Mecke, J. (1995). *Stochastic geometry and its applications*. 2nd. ed. Chichester; New York: Wiley.

Straub, C., Weinacker, H. & Koch, B. (2008). A fully automated procedure for delineation and classification of forest and non-forest vegetation based on full waveform laser scanner data. In: Chen, J., Jiang, J. & Peled, A. (eds) *ISPRS Archives – Volume XXXVII Part B8.* Bejing: ISPRS, pp. 1013-1020.

Tansey, K., Chambers, I., Anstee, A., Denniss, A. & Lamb, A. (2009). Object-oriented classification of very high resolution airborne imagery for the extraction of hedgerows and field margin cover in agricultural areas. *Applied Geography*, 29, pp. 145-157.

Taubenbock, H., Esch, T., Wurm, M., Roth, A. & Dech, S. (2010). Object-based feature extraction using high spatial resolution satellite data of urban areas. *Journal of Spatial Science,* 55(1), pp. 117-132.

Tewari, V.P., Sukumar, R., Kumar, R. & Gadow, K.v. (2014). Forest observational studies in India: Past developments and considerations for the future. *Forest Ecology and Management,* 316, pp. 32-46.

Thornton, M.W., Atkinson, P.M. & Holland, D.A. (2006). Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping. *International Journal of Remote Sensing*, 27, pp. 473-491.

Thornton, M.W., Atkinson, P.M. & Holland, D.A. (2007). A linearised pixel-swapping method for mapping rural linear land cover features from fine spatial resolution remotely sensed imagery. *Computers & Geosciences,* 33(10), pp. 1261-1272.

Tscheschel, A. & Stoyan, D. (2006). Statistical reconstruction of random point patterns. *Computational Statistics and Data Analysis,* 51, pp. 859-871.

UNCCD (1994). *United Nations Convention to Combat Desertification*. New York: UN.

UNFCCC (2008). *Kyoto Protocoll Reference Manual*. Bonn, Germany: UNFCCC.

Walker, J.S. & Briggs, J.M. (2007). An object-oriented approach to urban forest mapping in Phoenix. *Photogrammetric engineering and remote sensing,* 73(5), pp. 577-583.

Walton, J.T. (2008). Difficulties with estimating city-wide urban forest cover change from national, remotely-sensed tree canopy maps. *Urban Ecosystems,* 11, pp. 81-90.

World Bank *Data | The World Bank*. http://data.worldbank.org/ [30/04/2014].

Yeong, C.L.Y. & Torquato, S. (1998). Reconstructing random media. *Physical Review E,* 57(1), pp. 495-506.

Yoon, T.K., Park, C.W., Lee, S.J., Ko, S., Kim, K.N., Son, Y., Lee, K.H., Oh, S., Lee, W.K. & Son, Y. (2013). Allometric equations for estimating the aboveground volume of five common urban street tree species in Daegu, Korea. *Urban Forestry & Urban Greening,* 12(3), pp. 344-349.

Zarnoch, S.J. & Bechtold, W.A. (2000). Estimating mapped-plot forest attributes with ratios of means. *Canadian Journal of Forest Research,* 30, pp. 688-697.

Zhou, W. & Troy, A. (2008). An object-oriented approach for analysing and characterizing urban landscape at the parcel level. *International Journal of Remote Sensing,* 29, pp. 3119-3135.

Zhou, X., Brandle, J.R., Awada, T.N., Schoeneberger, M.M., Martin, D.L., Xin, Y. & Tang, Z. (2011). The use of forest-derived specific gravity for the conversion of volume to biomass for open-grown trees on agricultural land. *Biomass and Bioenergy,* 35(5), pp. 1721-1731.

Zhou, X., Schoeneberger, M.M., Brandle, J.R., Awada, T.N., Chu, J., Martin, D.L., Li, J., Li, Y. & Mize, C.W. (2014). Analyzing the Uncertainties in Use of Forest-Derived Biomass Equations for Open-Grown Trees in Agricultural Land. *Forest Science,* 61(1), pp. 144-161.

Zhou, X.H., Brandle, J.R., Schoeneberger, M.M. & Awada, T. (2007). Developing above-ground woody biomass equations for open-grown, multiple-stemmed tree species: Shelterbelt-grown Russian-olive. *Ecological Modelling,* 202(3-4), pp. 311-323.

Zomer, R.J., Coe, R., Place, F., van Noordwijk, M. & Xu, J.C. (2014). *Trees on farms: an update and reanalysis of agroforestry's global extent and socio-ecological characteristics*. Bogor, Indonesia: World Agroforestry Centre (ICRAF) Southeast Asian Regional Program.

Acknowledgements

First of all, I would like to thank my main supervisor, Göran Ståhl, for accepting me as a doctoral student at the department, and in particular for his kindness and patience and the always open door. Thank you Göran for everything, I learned a lot from you during the past four and half years.

I also owe a great debt of gratitude to Christoph Kleinn for waking the interest in the strange subject of counting trees (in and outside forests) and for the continuous commenting on my work and support.

Further, I want thank the Forest Faculty and the Department of Forest Resource Management at SLU for generously financing my project within the FONASO doctoral programme. I'm also deeply grateful to the people of Sweden, who made all this possible at the end.

Great thanks also to all my colleagues at the Department of Forest Resource Management for the wonderful administrative support and the friendly atmosphere that made it a pleasure to work here in Umeå.

The same can be said about the people at the Department of Geosciences and Natural Resource Management at Copenhagen University, where my family and I spent six months on a study visit. It was a wonderful time in Copenhagen. Here, special thanks go to my assistant supervisor Thomas Nord-Larsen for opening the door to Copenhagen and for all the support and guidance you gave me. My thanks also to Babro Haar for her friendliness and the always most welcome occasions of having a conversation in German. Further, I thank Vivian Kvist Johannsen for supporting our stay in Copenhagen and the PhD students at the department for having a nice time.

Also my deepest thanks to my assistant supervisor Arne Pommerening for all the good discussions and encouragements, you have been a great help.

Great thanks to all my co-authors for the support in writing the articles, and to Lena Hellstrandh for teaching me Swedish at the forest faculty in Umeå. I always enjoyed the two hours of language lessons we had each week together with fellow PhD students. Only by knowing Swedish, you really get to understand the country and its people.

The people of the Department of Geodesy and Geoinformation at Vienna University of Technology are gratefully acknowledged for granting a thesis license for their OPALS software that was used to process the ALS data from Paper III and Paper IV.

To my parents in law, Eva-Maria and Gunnar, for all the support they gave us during our time in Sweden. To my parents, for raising me up and for giving me the freedom to pursue my goals.

Finally, to my lovely wife Edda: Thank you for the beautiful drawing on the front cover and most importantly for our wonderful marriage and our gorgeous son Ragnar. I know that not everything turned out as expected for you; the more grateful I am to you for all your support and sacrifice.

Sebastian Schnell
Umeå, March 2015