PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND

Dissertations in Forestry and Natural Sciences

PETRI VARVIA

UNCERTAINTY QUANTIFICATION IN REMOTE SENSING OF FORESTS

UNIVERSITY OF
EASTERN FINLAND

PUBLICATIONS OF THE UNIVERSITY OF EASTERN FINLAND
DISSERTATIONS IN FORESTRY AND NATURAL SCIENCES

N:o 313

Petri Varvia

UNCERTAINTY QUANTIFICATION IN REMOTE SENSING OF FORESTS

ACADEMIC DISSERTATION

To be presented by the permission of the Faculty of Science and Forestry for public examination in the Auditorium TTA in Tietoteknia Building at the University of Eastern Finland, Kuopio, on August 31st, 2018, at 12 o'clock.

University of Eastern Finland
Department of Applied Physics
Kuopio 2018

Grano Oy
Jyväskylä, 2018
Editors: Pertti Pasanen, Matti Tedre,
Jukka Tuomela, and Matti Vornanen

Distribution:
University of Eastern Finland Library / Sales of publications
julkaisumyynti@uef.fi
http://www.uef.fi/kirjasto

ISBN: 978-952-61-2866-5 (print)
ISSNL: 1798-5668
ISSN: 1798-5668
ISBN: 978-952-61-2867-2 (pdf)
ISSNL: 1798-5668
ISSN: 1798-5676

Author's address:
University of Eastern Finland
Department of Applied Physics
P.O.Box 1627
FI-70211 Kuopio
Finland
email: petri.varvia@uef.fi

Supervisors:
Associate Professor Aku Seppänen
University of Eastern Finland
Department of Applied Physics
P.O.Box 1627
FI-70211 Kuopio
Finland
email: aku.seppanen@uef.fi

Docent Timo Lähivaara
University of Eastern Finland
Department of Applied Physics
P.O.Box 1627
FI-70211 Kuopio
Finland
email: timo.lahivaara@uef.fi

Professor Jari Kaipio
The University of Auckland
Department of Mathematics
Private Bag 92019
Auckland 1142
New Zealand
email: jari@math.auckland.ac.nz

Reviewers:
Professor Daniela Calvetti
Case Western Reserve University
Department of Mathematics,
Applied Mathematics and Statistics
10900 Euclid Avenue
Cleveland, OH 44106
USA
email: daniela.calvetti@case.edu

Assistant Professor Hans-Erik Andersen
USDA Forest Service
Pacific Northwest Research Station
P.O. Box 352100
Seattle, WA 98195-2100
USA
email: handersen@fs.fed.us

Opponent:
Docent Matti Mõttus
VTT Technical Research Centre of Finland
Department of Environmental Sciences
P.O. Box 1000
FI-02044 VTT Espoo
Finland
email: matti.mottus@gmail.com

Petri Varvia
Uncertainty quantification in remote sensing of forests
Kuopio: University of Eastern Finland, 2018
Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences

ABSTRACT

The aim of this thesis is to develop estimate uncertainty quantification methods for two different remote sensing problems: 1) Estimation of forest inventory attributes, such as stem volume, from airborne laser scanning (ALS) data. 2) Estimation of forest canopy leaf area index (LAI) from hyperspectral satellite images. The proposed approaches for both problems are based on Bayesian inference. In the Bayesian framework, both the measurements and the unknown variables are modeled as random variables. Prior distributions are constructed for the unknowns and are then updated using information gained from the measurements. The result of the Bayesian inference is a posterior density, that describes the information on the unknown variables.

For the forest attribute estimation from the ALS data, we present an approach to quantify the plot-level uncertainty in species-specific growing stock volume and other stand attributes within the so-called area-based approach. The results show that the proposed approach performs well in quantifying the estimate uncertainty and produces optimal interval estimates for species-specific volumes when sufficient training data is available. Also the point estimate accuracy is competitive with current state-of-the-art methods. We also demonstrate how the quantified uncertainties of the stand attributes can be utilized to determine the uncertainty in classification done using the estimated stand attributes.

In the estimation of canopy LAI, we present a method based on Bayesian inversion of a physically-based forest reflectance model. Forest reflectance model contain multiple other unknown variables in addition to LAI. In the proposed approach, these unknown parameters are estimated simultaneously with LAI from hyperspectral satellite data. We study the effects of reflectance model parameter uncertainties on LAI estimates in a simulation study. The results show that in the presence of unknown parameters, the Bayesian LAI estimates which account for the model uncertainties outperform the conventional estimates that are based on biased model parameters. Moreover, the results demonstrate that the Bayesian inference can also provide feasible measures for the uncertainty of the estimated LAI. Furthermore, the feasibility of the Bayesian approach is tested using hyperspectral EO-1 Hyperion data in a heterogenous boreal forest area at multiple dates over the growing season. The results show that the Bayesian inversion approach is significantly better than using a comparable spectral vegetation index regression for LAI estimation.

Universal Decimal Classification: *519.226, 528.8, 630*58*

Library of Congress Subject Headings: *Remote sensing; Forests and forestry; Forest biomass; Tree trunks; Forest canopies; Leaf area index; Aerial surveys in forestry; Lasers; Scanning systems; Artificial satellites in forestry; Hyperspectral imaging; Uncertainty; Estimation theory; Bayesian statistical decision theory*

Yleinen suomalainen asiasanasto: *kaukokartoitus; metsät; puusto; puut; laserkeilaus; spektrikuvaus; satelliittikuvaus; satelliittikuvat; epävarmuus; tarkkuus; estimointi; bayesilainen menetelmä*

ACKNOWLEDGEMENTS

This work was carried out at the Department of Applied Physics at the University of Eastern Finland, Kuopio during 2013–2018.

First, I would like to thank my principal supervisor Associate Professor Aku Seppänen for his support, guidance, and the tireless endeavor to improve my scientific writing skills. I would like to thank my second supervisor Docent Timo Lähivaara for his guidance, especially in practical matters. I would also like to thank my supervisor Professor Jari Kaipio. I wish to thank all my supervisors for giving me considerable creative freedom in this thesis project, and for the many interesting conversations over the years.

I would like to thank the official pre-examiners of this thesis, Professor Daniela Calvetti and Assistant Professor Hans-Erik Andersen, for their highly positive and helpful comments.

I wish to thank all my co-authors, Assistant Professor Miina Rautiainen, Associate Professor Petteri Packalen, Professor Matti Maltamo, and Professor Timo Tokola, for their contribution to this thesis. Interdisciplinary work like this would not have been possible without your expertise. I also wish to thank all the members of the Inverse Problems group at Kuopio, and by extension the Finnish inverse problems community.

Finally, I would like to thank my mother, Anne Pulkkinen, and my relatives for their support. I would also like thank my friends, especially Suvi Kummu, Laura Tiainen, and Tuomas Suihkonen, for their surprising tolerance to lengthy, incomprehensible monologues.

This study was supported by the University of Eastern Finland spearhead project "Multiscale geospatial analysis of forest ecosystems", the doctoral school of the University of Eastern Finland), the Finnish Cultural Foundation, North Savo Regional fund, and the Finnish Centre of Excellence of Inverse Problems Research 2012-2017.

Kuopio, August 8, 2018

Petri Varvia

LIST OF PUBLICATIONS

This thesis consists of the present review of the author's work in the field of remote sensing and the following selection of the author's publications:

I P. Varvia, T. Lähivaara, M. Maltamo, P. Packalen, T. Tokola, and A. Seppänen, "Uncertainty quantification in ALS-based species-specific growing stock volume estimation," *IEEE Transactions on Geoscience and Remote Sensing* **55**(3) 1671–1681 (2017).

II P. Varvia, M. Rautiainen, and A. Seppänen, "Modeling uncertainties in estimation of canopy LAI from hyperspectral remote sensing data – A Bayesian approach," *Journal of Quantitative Spectroscopy and Radiative Transfer* **191**, 19–29 (2017).

III P. Varvia, M. Rautiainen, and A. Seppänen, "Bayesian estimation of seasonal course of canopy leaf area index from hyperspectral satellite data," *Journal of Quantitative Spectroscopy and Radiative Transfer* **208**, 19–28 (2018).

Throughout the overview, these papers will be referred to by Roman numerals.

AUTHOR'S CONTRIBUTION

The publications selected in this dissertation are original research papers on remote sensing of forests. The author implemented all the numerical methods and did the computations in Matlab. The author of this thesis was the principal writer in all the publications.

TABLE OF CONTENTS

ABBREVIATIONS

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
ABA	area-based approach
ALS	airborne laser scanning
bias%	relative bias
BRF	bidirectional reflectance factor
CI	credible interval
CI%	credible interval coverage
CM	conditional mean
DRAM	delayed rejection adaptive Metropolis
HDRF	hemispherical-directional reflectance factor
kNN	*k* nearest neighbor
LAI	leaf area index
LiDAR	light detection and ranging
LOO	leave-one-out
LUT	look-up table
MAP	maximum a posteriori
MCMC	Markov chain Monte Carlo
ML	maximum likelihood
MSN	most similar neighbor
RMSE	root mean square error
RMSE%	relative root mean square error
SRWI	simple ratio water index
VI	vegetation index

NOMENCLATURE

GENERAL

$\propto$	directly proportional to
a	scalar
$\mathbf{a}$	vector
$\mathbf{A}$	matrix
arg max	maximum point
arg min	minimum point
$\mathbb{R}^n$	n-dimensional real space
$\pi(\cdot)$	probability density function
$\boldsymbol{\mu}$	expected value
$\boldsymbol{\Gamma}$	covariance matrix
$\mathbf{R}$	correlation matrix
$\boldsymbol{\theta}$	vector of unknowns
$\mathbf{z}$	measurement vector
$\mathbf{e}$	error term

RELATED TO ALS-BASED FOREST INVENTORY

h	tree height
d	stem diameter
N	stem number
BA	basal area
v	stem volume
$\mathbf{z}_t, \boldsymbol{\theta}_t$	$\mathbf{z}$ and $\boldsymbol{\theta}$ of a training plot
$(\mathbf{Z}_t, \boldsymbol{\Theta}_t)$	set of training data
$\boldsymbol{\phi}(\boldsymbol{\theta})$	vector of basis functions
$\boldsymbol{\Phi}$	design matrix for the linear regression
$\hat{\mathbf{A}}$	model matrix of the fitted linear model
$\hat{\boldsymbol{\mu}}$	sample mean
$\hat{\boldsymbol{\Gamma}}$	sample covariance
$\boldsymbol{\Sigma}$	additional error variance matrix
$p_{\text{dominance},i}$	dominance probability for the i'th species

RELATED TO LAI ESTIMATION

λ	wavelength
LAI_{eff}	effective leaf area index
β	shoot clumping factor
ω_L	leaf single scattering albedo
ρ_g	understory BRF
p	photon recollision probability
i_c	canopy interceptance
t_c	canopy transmittance
t_d	diffuse canopy transmittance
$S(\lambda; \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\omega}}_L)$	spline approximation for $\boldsymbol{\omega}_L$ with the nodes $\tilde{\boldsymbol{\lambda}}$
$\tilde{\boldsymbol{\omega}}_L$	ω_L on the spline approximation nodes
$\tilde{\boldsymbol{\rho}}_g$	ρ_g on the spline approximation nodes

1 Introduction

Roughly third of the global land surface area is covered by forests [1]. Forests contain the majority of terrestrial plant biomass and carbon, and also the majority of terrestrial animal species. In addition to this ecological importance, forests have enormous economical and cultural value through, for example, timber production and recreation. Due to the ongoing global warming and deforestation, forest ecosystems are under change. Effective management and conservation of forest resources in these changing conditions requires timely information on the state of the forests, which due to the vast geographical span involved can be realistically acquired only by using remote sensing, that is, measurements and derived data products from airborne or spaceborne sensors.

The remote sensed data often contain only indirect information on the forest attributes of interest. Computational and statistical methods are then required to estimate the attributes of interest based on these data. Commonly used methods to produce estimates from remote sensed data give only point estimates for the quantity. Uncertainty of these point estimates is then usually considered only in classical statistical sense, for example, how reliable is the calculated mean of the forest attribute over a large area? Less attention has been directed to the problem of producing feasible error intervals corresponding to individual point estimates, i.e. Bayesian uncertainty quantification. When the remote sensing estimates are further used, for example, as input parameters in climate or biological models, or for decision making or risk assessment, such information on the uncertainty of the estimates would be crucial (e.g. [2, 3]).

The aim of this thesis is to develop Bayesian estimation and uncertainty quantification methods specifically for two different remote sensing problems:

1. Estimation of forest inventory attributes, such as stem volume, from airborne laser scanning (ALS) data.

2. Estimation of forest canopy leaf area index (LAI) from hyperspectral satellite images.

In the Bayesian setting, both the measurements (e.g. satellite measured reflectance) and the unknown variables of interest are modeled as random variables (e.g. [4–6]). For the unknowns, prior distributions are constructed; they describe the *a priori* known information on the variables. The prior formulation thus also inherently models the *a priori* uncertainty of these variables, which is perhaps the most significant advantage of Bayesian methods. The prior distribution is then updated using information gained from the measurements. This is done through a likelihood function that includes a model that connects the unknown variables to the measurement and information on the uncertainty of the measurement. The result of Bayesian inference is a posterior distribution, that describes the information on the unknown variables given the prior formulation and the measured data. From the posterior density, point estimates and uncertainty metrics are then calculated.

1.1 ESTIMATION OF STAND ATTRIBUTES FROM AIRBORNE LASER SCANNING DATA

Airborne laser scanning (ALS) is a LiDAR (light detection and ranging) -based remote sensing modality that generates a point cloud that captures the three-dimensional structure of the target (e.g. [7,8]). Because of this, ALS is efficient in capturing forest height and attributes that depend on the height, such as stem volume. ALS is used for forest inventories, forest monitoring, and forest ecological studies, among other applications. Schematic of the measurement setup of ALS is shown in Figure 1.1.

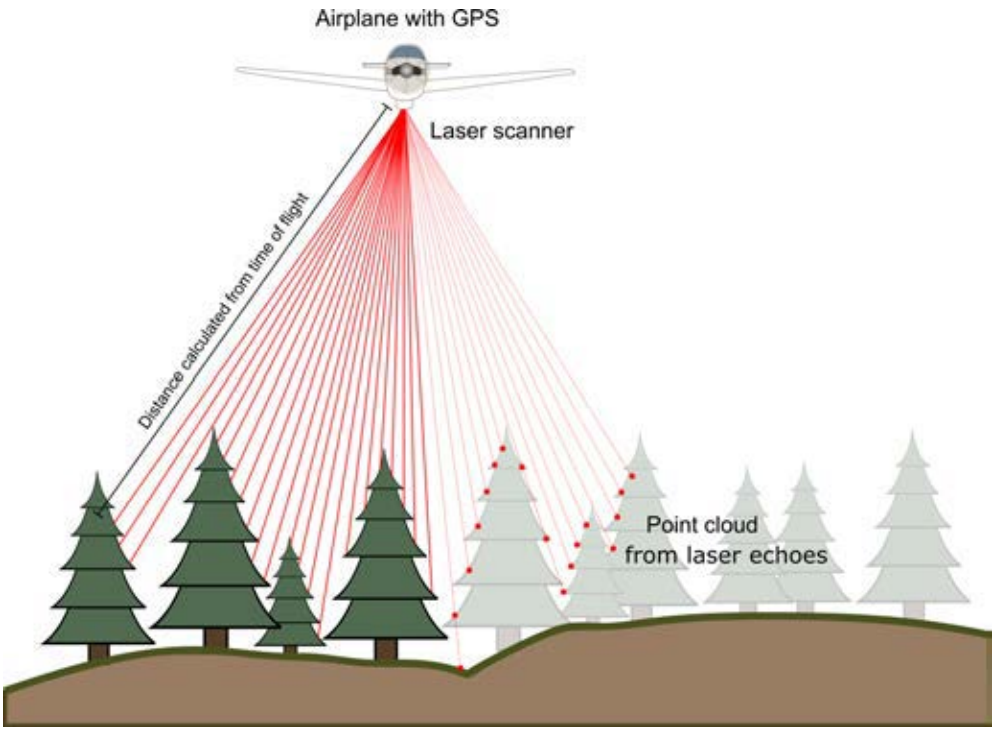

Figure 1.1: Simplified schematic of the ALS measurement setup.

Methods for ALS-based forest inventories [9, 10] can be generally divided into two groups: the area-based approach (ABA) (e.g. [11–15]) and single-tree detection (e.g. [16–21]). In ABA, stand attributes are estimated from the ALS point cloud by first calculating various metrics from the height distribution of the laser returns within a small forest segment/plot and then by formulating a regression model between these metrics and field-measured stand attributes using a number of training plots. The stand attributes for the remaining segments are predicted using the trained regression model. Usually metrics computed from aerial images are used in addition to ALS height distribution metrics. In single-tree detection, on the other hand, individual tree crowns are detected from high-density ALS data and stand attributes are subsequently predicted usually by allometric relationships between the stand attributes and crown geometry. Some hybrid approaches have also been proposed that have combined characteristics of both ABA and single-tee detection. Publication **I** concentrates on the area-based approach.

In the area based approach, the current estimation methods include linear regression (e.g. [11, 12]), sparse Bayesian estimation [22], and non-parametric regression methods, such as k nearest neighbor (kNN) algorithm (e.g. [13, 23, 24]), and artificial neural networks [25, 26]. These existing methods have usually a good accuracy, but in most cases do not produce uncertainty estimates. So far, a few methods to predict plot/cell level variance have been proposed [22, 27, 28]. At stand level, studies on model-based variance estimation have been published, e.g. [29, 30].

In the publication **I**, a Bayesian inference approach for estimation of stand attributes within the ABA framework was proposed. The approach was tested in estimation of species-specific stem volumes. The method produces uncertainty estimates for the attributes and can be used robustly with a large number of ABA metrics. Furthermore, the parameter uncertainties were used to evaluate uncertainty in classification by dominant species. The method and the results are summarized in Chapter 3.

1.2 ESTIMATION OF CANOPY LEAF AREA INDEX FROM HYPERSPECTRAL DATA

Leaf area index (LAI) is usually defined as half of the total leaf/needle area per unit ground surface area [31, 32]. Throughout this thesis, LAI specifically refers to the LAI of the forest canopy. Because leaf area, and thus LAI, is directly related to light interception and biological processes, principally gas exchange, it is an important biophysical variable. Accurate estimates of LAI are required to model vegetation carbon and water cycles and are thus by proxy important for climate models [33, 34].

Various methods to estimate forest canopy LAI from multispectral (a few wide spectral bands) and hyperspectral (many narrow spectral bands) satellite images have been proposed over the years. The rationale is that the radiation scattered from the forest canopy carries information on the canopy parameters, including LAI (Figure 1.2). A classical empirical approach is to produce so-called spectral vegetation indices (VI) from the satellite measurement that show strong correlation with LAI. The VI can be, for example, a ratio of two spectral bands, or a more complex transformation. Using training data, a regression model for LAI given the VI is then formed and used to estimate LAI. The main problem of VI regressions is their strong site, species, and time specificity (e.g. [35–37]).

Approaches based on forest reflectance model inversion are seen as a solution to the specificity problems of empirical VI regression. Forest reflectance models are mathematical models that describe the top of the canopy reflectance of the forest based on geometric and optical properties of the forest. Forest reflectance models, however, contain numerous other variables beside LAI, which makes the model inversion difficult: an ill-posed inverse problem [38, 39]. The reflectance model inversion based approaches can be roughly divided to two main classes: 1) Methods based on generating a large simulated training set and then using VI regression (e.g. [40, 41]) or look up table (LUT) based methods (e.g. [42,43]) to predict LAI using the synthetic data set. 2) Direct inversion of the reflectance model by using, for example, classical regularization techniques and numeric optimization (e.g. [44, 45]). All these methods have drawbacks. Synthetic VI regression always discards part of the information in the data in reducing it to a single VI. In LUTs, the main problem is to find balance between generality, number of variables and manageable LUT size. Optimization

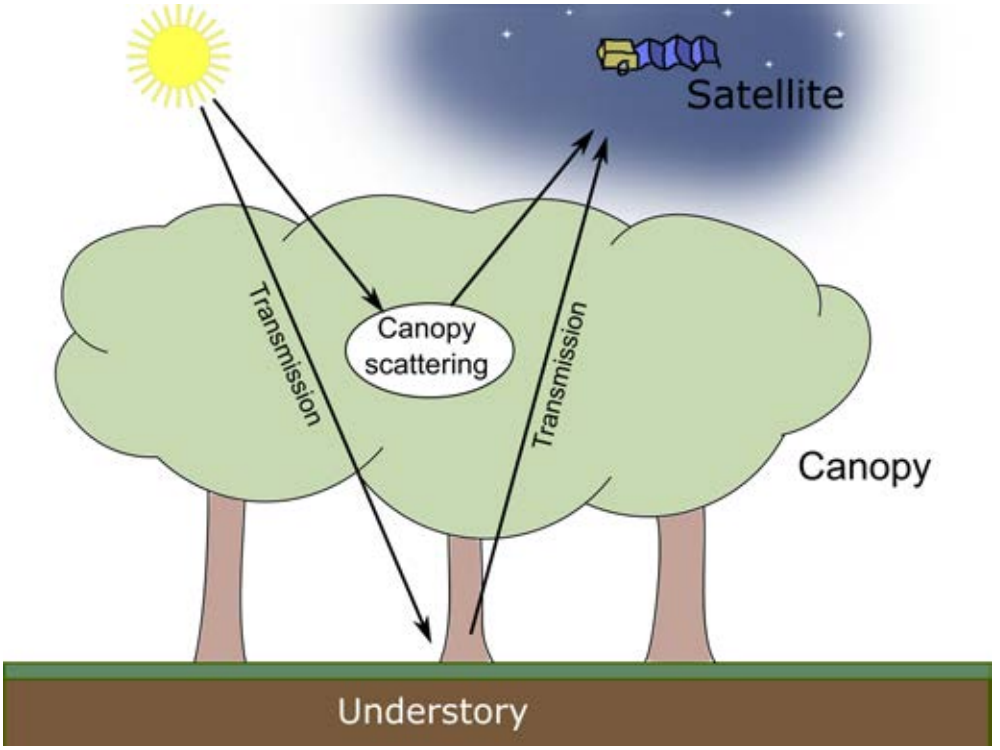

Figure 1.2: Schematic picture of the basic setup of optical satellite remote sensing of a forest canopy.

based reflectance model inversion methods suffer from the strong nonlinearity of the reflectance models, and from the fact that the problem often has no unique solution.

In this work, we propose a Bayesian approach to forest reflectance model inversion using hyperspectral satellite data. The proposed approach is based on constructing informative prior distributions for the reflectance model parameters and exploration of the resulting posterior distribution using a Monte Carlo method. Bayesian approach has been previously used in remote sensing of forest structural parameters from multispectral MODIS data by Zhang *et al.* [46], where prior model consisted only of constraints for the model unknowns.

The testing of the Bayesian approach in this thesis was done in two parts. First, in the publication **II**, the approach was tested using simulated data. In a simulation study, all the model parameters can be controlled and thus their effect is more easily quantified. The Bayesian approach was compared to VI regression and an optimization based method, and the effect of informative versus uninformative prior model was also evaluated. In the reference optimization based method, we show that fixing the unknown model (nuisance) parameters to their best-guess values can introduce large errors. The Bayesian approach was then tested using real EO-1 Hyperion data over a boreal forest area in the publication **III**. There, the Bayesian approach was compared to VI regression using both real and simulated training data. In both phases the quality of the estimate uncertainty metrics produced by the Bayesian approach was also examined. The reflectance model inversion and the results are summarized in Chapter 4.

2 Bayesian inference and uncertainty quantification

In this chapter, the theory of Bayesian inference is shortly reviewed. For consistency and simplicity, unified variable notation is introduced here. The notation in the following chapters thus differs somewhat from the notations used in the publications **I – III**. Vector-valued variables are bolded, and matrices bolded and capitalized. Let $\boldsymbol{\theta} \in \mathbb{R}^n$ be the vector of variables of interest, that is, the variables that are to be estimated, and let $\mathbf{z} \in \mathbb{R}^m$ be the vector of measurements. Both $\boldsymbol{\theta}$ and $\mathbf{z}$ are modeled here as random variables (e.g. [4]).

2.1 BAYESIAN INFERENCE

In a general setting, the variables $\boldsymbol{\theta}$ and the measurements $\mathbf{z}$ are connected by an observation model of the form

$$\mathbf{z} = f(\boldsymbol{\theta}, \boldsymbol{\xi}, \mathbf{e}), \tag{2.1}$$

where $f(\,\cdot\,)$ is a function, $\boldsymbol{\xi}$ is a vector of confounding model parameters, and $\mathbf{e}$ is an error term that is the discrepancy between the model output and reality. Usually, $\mathbf{e}$ is modeled as an additive error term and $\boldsymbol{\xi}$ is assumed to be known, resulting in the observation model

$$\mathbf{z} = f(\boldsymbol{\theta}) + \mathbf{e}. \tag{2.2}$$

A useful property of the form (2.2) is that the conditional density $\pi(\mathbf{z}|\boldsymbol{\theta})$ of $\mathbf{z}$ given $\boldsymbol{\theta}$ can be formulated as [4]

$$\pi(\mathbf{z}|\boldsymbol{\theta}) = \pi_{\mathbf{e}}(\mathbf{z} - f(\boldsymbol{\theta})), \tag{2.3}$$

where $\pi_{\mathbf{e}}(\,\cdot\,)$ is the probability density function of the error $\mathbf{e}$. The density $\pi(\mathbf{z}|\boldsymbol{\theta})$ is usually called the likelihood function and describes the distribution of $\mathbf{z}$ given $\boldsymbol{\theta}$. Error $\mathbf{e}$ is often modeled as a Gaussian random variable, in which case the likelihood function is

$$\pi(\mathbf{z}|\boldsymbol{\theta}) \propto \exp\left(-\frac{1}{2}(\mathbf{z} - f(\boldsymbol{\theta}) - \boldsymbol{\mu}_{\mathbf{e}})^T \boldsymbol{\Gamma}_{\mathbf{e}}^{-1} (\mathbf{z} - f(\boldsymbol{\theta}) - \boldsymbol{\mu}_{\mathbf{e}})\right), \tag{2.4}$$

where $\boldsymbol{\mu}_{\mathbf{e}}$ is the expected value and $\boldsymbol{\Gamma}_{\mathbf{e}}$ the covariance matrix of $\mathbf{e}$. When the measurement $\mathbf{z}_{\text{meas.}}$ is performed, the likelihood function $\pi(\mathbf{z}_{\text{meas.}}|\boldsymbol{\theta})$ is a function of $\boldsymbol{\theta}$, with $\mathbf{z}$ fixed to $\mathbf{z}_{\text{meas.}}$. In the rest of this thesis, $\mathbf{z}$ denotes this measured $\mathbf{z}_{\text{meas.}}$. If the statistics of $\mathbf{e}$ are not well known (as is often the case), it is usually approximated to be zero mean and uncorrelated, as in the articles **II** and **III**. However, the statistics of $\mathbf{e}$ can be estimated based on real (as in the publication **I**) or synthetic training data (see e.g. [4,47]) when such data is available.

Given the measured vector $\mathbf{z}$, the Bayesian inference problem is to formulate the conditional density $\pi(\boldsymbol{\theta}|\mathbf{z})$, which describes the distribution of $\boldsymbol{\theta}$ given the measured $\mathbf{z}$ (see e.g. [4,5]). This can be accomplished by using the Bayes' rule

$$\pi(\boldsymbol{\theta}|\mathbf{z}) = \frac{\pi(\mathbf{z}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{z})} \propto \pi(\mathbf{z}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}), \tag{2.5}$$

where $\pi(\mathbf{z}|\boldsymbol{\theta})$ is the previously introduced likelihood function, $\pi(\boldsymbol{\theta})$ is the so-called prior density, and $\pi(\mathbf{z})$ functions as a normalizing constant and can be usually ignored. The conditional density $\pi(\boldsymbol{\theta}|\mathbf{z})$ is referred to as the posterior density. The prior density $\pi(\boldsymbol{\theta})$ contains the *a priori* information on $\boldsymbol{\theta}$, that is, the inferrer's preconception on what values of $\boldsymbol{\theta}$ are probable before the information gained from the measurement $\mathbf{z}$ is applied. This prior state is then updated using the information gained from the measurement (the likelihood function) to produce the *a posteriori* state of information (the posterior density).

2.1.1 Prior modeling

The prior probability density $\pi(\boldsymbol{\theta})$ is a crucial part of the Bayesian framework. The goal in prior modeling is to formulate a prior density that accurately describes the available information on the variable, including its natural variation, constraints on feasible values, and structure.

In this thesis, two common families of prior densities are used: uniform priors and Gaussian priors. In uniform priors, the prior density is the uniform distribution

$$\pi(\boldsymbol{\theta}) \propto \begin{cases} 1 & \boldsymbol{\theta} \in \mathcal{G} \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

where $\mathcal{G}$ is the domain to which $\boldsymbol{\theta}$ is constrained. Usually $\mathcal{G}$ is a box, that is, a Cartesian product of intervals $[a_i, b_i]$. Such box constraint was used in the publications **II** and **III**. In the publication **I**, the so-called positivity or non-negativity constraint was used, in which $\mathcal{G}$ is $\mathbb{R}^n_+$, i.e. the positive part of the real coordinate space. It should be noted, that while the positivity constraint does not result in a proper prior density, the resulting posterior density is a proper probability density, if the likelihood function is non-singular.

In Gaussian priors, the prior distribution is the multivariate normal distribution, with the prior probability density

$$\pi(\boldsymbol{\theta}) \propto \exp\left(-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\mu}_{\boldsymbol{\theta}})^T \boldsymbol{\Gamma}_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu}_{\boldsymbol{\theta}})\right), \tag{2.7}$$

where $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ is the prior expectation and $\boldsymbol{\Gamma}_{\boldsymbol{\theta}}$ is the prior covariance matrix. The prior (hyper)parameters $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ and $\boldsymbol{\Gamma}_{\boldsymbol{\theta}}$ can be based on, for example, training data or previous knowledge of the expected value, range, and structure of $\boldsymbol{\theta}$. The prior covariance matrix can be further decomposed to

$$\boldsymbol{\Gamma}_{\boldsymbol{\theta}} = \mathbf{S}_{\boldsymbol{\theta}} \mathbf{R}_{\boldsymbol{\theta}} \mathbf{S}_{\boldsymbol{\theta}}, \tag{2.8}$$

where $\mathbf{R}_{\boldsymbol{\theta}}$ is the correlation matrix (i.e. the matrix of Pearson correlation coefficients) and $\mathbf{S}_{\boldsymbol{\theta}}$ is a diagonal matrix containing the prior standard deviations of $\boldsymbol{\theta}$ on its main diagonal. Using this decomposition, prior correlation structure and prior variances can be prescribed separately, as was done in the articles **II** and **III**.

2.2 ESTIMATES AND COMPUTATION

Once the posterior density $\pi(\boldsymbol{\theta}|\mathbf{z})$ is formed using the prior and the likelihood, descriptive point and interval statistics are usually computed. The choice of statistic

depends on what is wanted from the estimate, on the shape of the posterior density, and on what is computationally feasible.

The classical Bayesian point estimate is the posterior conditional expectation/mean:

$$\boldsymbol{\theta}_{\mathrm{CM}} = \int_{\mathbb{R}^n} \boldsymbol{\theta}\pi(\boldsymbol{\theta}|\mathbf{z})\mathrm{d}\boldsymbol{\theta}. \tag{2.9}$$

The posterior conditional mean (CM) estimate is unbiased and has the minimum variance of estimate error, if the statistical modeling is accurate [48]. The drawbacks of the CM estimate are the often high computational cost due to the required integration, and the occasional poor behavior on multimodal or strongly skewed posterior distributions.

Another, perhaps even more commonly used, point estimate is the posterior mode, or maximum a posteriori (MAP) estimate:

$$\boldsymbol{\theta}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{\theta}}\{\pi(\boldsymbol{\theta}|\mathbf{z})\}. \tag{2.10}$$

The most desirable property of the MAP estimate is that it can be computed by solving an optimization problem, which in most cases is substantially faster than integrating over the posterior. MAP can also produce subjectively better estimates in problems where, for example, sparse solutions are wanted.

In theory, the posterior (geometric) median [49] can also be calculated, thus completing the classical triumvirate of central tendency measures:

$$\boldsymbol{\theta}_{\mathrm{MD}} = \arg\min_{\tilde{\boldsymbol{\theta}}}\left\{\int_{\mathbb{R}^n} \|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}\|\,\pi(\boldsymbol{\theta}|\mathbf{z})\mathrm{d}\boldsymbol{\theta}\right\}. \tag{2.11}$$

However, due to the substantial computational cost, multivariate median estimates are rarely used. The possible benefits of using posterior median instead of posterior mean include better robustness when the posterior density is heavy tailed [50].

2.2.1 Estimates of dispersion

In addition to the point estimate that measures the central tendency, suitable estimate of dispersion is needed. The dispersion describes the uncertainty of the estimate. Perhaps the simplest dispersion estimate is the posterior covariance $\boldsymbol{\Gamma}_{\boldsymbol{\theta}|\mathbf{z}}$, from which the posterior variances (i.e. diagonal elements) $\sigma^2_{\theta_i|\mathbf{z}}$ for each single variable θ_i can be extracted and the corresponding standard deviations $\sigma_{\theta_i|\mathbf{z}}$ calculated.

However, if the posterior density is non-Gaussian, which is often the case, standard deviation is not a good uncertainty estimate, because for an arbitrary posterior density the probability of e.g. the 1σ interval $[\theta_{\mathrm{CM}} - \sigma_{\theta|\mathbf{z}}, \theta_{\mathrm{CM}} + \sigma_{\theta|\mathbf{z}}]$ is not necessarily readily available from the literature, as it is for the normal distribution. Furthermore, if the posterior density is skewed, the uncertainty is direction dependent, while σ interval assumes symmetric dispersion.

In the above cases, it is more sensible to compute a credible interval. For example a 95% credible interval ($\mathrm{CI}_{95\%}$) is an interval $[a, b]$ which satisfies the equation: [51]

$$\int_a^b \pi(\theta_i|\mathbf{z})d\theta_i = 0.95, \tag{2.12}$$

where $\pi(\theta_i|\mathbf{z})$ is the posterior marginal density

$$\pi(\theta_i|\mathbf{z}) = \int_{\mathbb{R}^{n-1}} \pi(\boldsymbol{\theta}|\mathbf{z})\mathrm{d}\boldsymbol{\theta}_{\{j\,:\,j\neq i\}}. \tag{2.13}$$

Credible intervals with other probabilities (e.g. 90%, 50%) can be calculated by substituting the desired probability on the right hand side of equation (2.12).

The equation (2.12) has no unique solution. Instead, a and b must be chosen according to a suitable rule. Two commonly used rules are: 1) choosing the narrowest interval, and 2) choosing the interval such that the remaining probability below a is equal to the probability above b (i.e. equal tails). In this thesis, mostly the first rule is used; the second rule is used in some visualizations in the articles **II** and **III**.

The posterior marginal densities (2.13) are in themselves useful tools for posterior density visualization, in both one-dimensional form (as in the equation) and for two variables, in which case it can be used to study the posterior interaction of the variables. In addition to one-dimensional credible intervals, multivariate credible regions could be also computed (e.g. [52]).

2.2.2 Markov chain Monte Carlo

Computation of posterior mean, median, marginal densities and credible intervals requires integration of the posterior density $\pi(\boldsymbol{\theta}|\mathbf{z})$. While in some special cases the integration can be accomplished analytically, in general the integration has to be done numerically. One widely used family of numerical integration algorithms in this setting is the Metropolis type algorithms, a subclass of Markov chain Monte Carlo (MCMC) methods (e.g. [53]).

In MCMC methods, a specifically constructed Markov chain is used to generate a sequence of N samples which ultimately statistically converges to the underlying posterior distribution. The computed finite sequence of these MCMC samples $\boldsymbol{\theta}^{(i)},\ i = 1, \ldots, N$ is used to approximate integrals:

$$\int_{\Omega} g(\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\mathbf{z})\mathrm{d}\boldsymbol{\theta} \approx \frac{1}{N}\sum_{i=1}^{N} g(\boldsymbol{\theta}^{(i)}),\ \ \forall i\ :\ \boldsymbol{\theta}^{(i)} \in \Omega, \tag{2.14}$$

where $g(\,\cdot\,)$ is an arbitrary function and Ω the integration domain. For example, the posterior mean estimate (2.9) is computed using the MCMC samples as

$$\boldsymbol{\theta}_{\mathrm{CM}} \approx \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\theta}^{(i)}. \tag{2.15}$$

To evaluate a 95% credible interval using the Monte Carlo samples requires finding the points a and b such that the number of samples $\hat{N}$ within the interval $[a, b]$ equals to $0.95 \cdot N$. For example, equal tails CI can be easily calculated by sorting the samples in an ascending order and choosing the $0.025 \cdot N$'th sample as a and $0.975 \cdot N$'th sample as b. The minimum width CI requires solving of an optimization problem.

In this thesis, the delayed rejection adaptive Metropolis (DRAM) [54] is used. DRAM is an extension of the classical Metropolis algorithm that includes additional components that make it more robust to poor choice of the sampling density. For description of the algorithm, see the publication **II**.

3 Uncertainty quantification in area-based approach

In this chapter, the methodology and results of the publication **I** are reviewed and summarized. The publication **I** concentrated on the prediction of species-specific growing stock volume from ALS data in the area-based framework. This chapter also introduces additional results, where multiple species-specific stand attributes (tree height, stem diameter, stem number, basal area, and stem volume) are estimated simultaneously.

3.1 PROBLEM SUMMARY

The mathematical formulation of the area-based approach (ABA) problem is as follows: Given a set of known, field-measured $n_{\boldsymbol{\theta}}$ stand attributes on n_t segments (training plots)

$$\boldsymbol{\Theta}_t = \begin{bmatrix} \boldsymbol{\theta}_t^{(1)} & \boldsymbol{\theta}_t^{(2)} & \dots & \boldsymbol{\theta}_t^{(n_t)} \end{bmatrix}^T \in \mathbb{R}^{n_t \times n_{\boldsymbol{\theta}}} \tag{3.1}$$

and the corresponding known, measured $n_{\mathbf{z}}$ LiDAR and aerial image predictors

$$\mathbf{Z}_t = \begin{bmatrix} \mathbf{z}_t^{(1)} & \mathbf{z}_t^{(2)} & \dots & \mathbf{z}_t^{(n_t)} \end{bmatrix}^T \in \mathbb{R}^{n_t \times n_{\mathbf{z}}}, \tag{3.2}$$

how to predict the unknown $\boldsymbol{\theta} \in \mathbb{R}^{n_{\boldsymbol{\theta}}}$ given a new measured predictor vector $\mathbf{z} \in \mathbb{R}^{n_{\mathbf{z}}}$?

The data set used in the publication **I** and in the extended results of this chapter consists of $n_t = 493$ field-measured plots and the corresponding predictors formed from the LiDAR point cloud and orthorectified multispectral aerial images. Five different stand attributes were measured (tree height h, stem diameter d, stem number N, basal area BA, and stem volume v) for three tree species groupings (pine, spruce, and deciduous), thus $n_{\boldsymbol{\theta}} = 15$. A total of $n_{\mathbf{z}} = 77$ predictors were used in the estimation of the forest attributes. For detailed description of the data, see the publication **I**.

3.2 BAYESIAN INVERSION OF AN EMPIRICAL MODEL

The training data $(\mathbf{Z}_t, \boldsymbol{\Theta}_t)$ is used to construct an empirical model. In the publication **I**, an empirical linear model was used

$$\mathbf{z} = \mathbf{A}\boldsymbol{\phi}(\boldsymbol{\theta}) + \mathbf{e}, \tag{3.3}$$

where $\mathbf{A} \in \mathbb{R}^{n_{\mathbf{z}} \times n_{\phi}}$ is a matrix, $\mathbf{e}$ a Gaussian additive error term and $\boldsymbol{\phi}(\boldsymbol{\theta})$ a vector of basis functions. In the publication **I**, the vector of basis functions was

$$\boldsymbol{\phi}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{v} \\ \frac{1}{\sum_{i=1}^{3} \mathbf{v}_i} \mathbf{v} \end{bmatrix} \in \mathbb{R}^6, \tag{3.4}$$

that is, a vector containing the stem volumes $\mathbf{v}$ and the relative stem volumes $\frac{1}{\sum_{i=1}^{3}\mathbf{v}_i}\mathbf{v}$. In the extended results, the basis functions are

$$\boldsymbol{\phi}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{h} \\ \mathbf{d} \\ \mathbf{N} \\ \mathbf{BA} \\ \mathbf{v} \\ \frac{1}{\sum_{i=1}^{3}\mathbf{v}_i}\mathbf{v} \end{bmatrix} \in \mathbb{R}^{18}, \tag{3.5}$$

where $\mathbf{h}$, $\mathbf{d}$, $\mathbf{N}$, and $\mathbf{BA}$ are vectors of species-specific tree height, stem diameter, stem number, and basal area, respectively.

The model (3.3) is fitted to the training data in least squares sense. First, the design matrix $\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\phi}(\boldsymbol{\theta}_t^{(1)}) & \boldsymbol{\phi}(\boldsymbol{\theta}_t^{(2)}) & \dots & \boldsymbol{\phi}(\boldsymbol{\theta}_t^{(n_t)}) \end{bmatrix}^T$ is formed. Then the least squares estimate for $\hat{\mathbf{A}}$ is

$$\hat{\mathbf{A}} = \mathbf{Z}_t^T \boldsymbol{\Phi}(\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}. \tag{3.6}$$

With $\hat{\mathbf{A}}$ now fixed, the residuals $\mathbf{e}_t^{(i)} = \mathbf{z}_t^{(i)} - \hat{\mathbf{A}}\boldsymbol{\phi}(\boldsymbol{\theta}_t^{(i)})$, $i = 1, \dots, n_t$ are computed and the $\boldsymbol{\theta}$-conditional statistics of $\mathbf{e}$ (the mean $\boldsymbol{\mu}_{\mathbf{e}|\boldsymbol{\theta}}$ and covariance $\boldsymbol{\Gamma}_{\mathbf{e}|\boldsymbol{\theta}}$) are approximated using sample statistics of $\mathbf{e}_t^{(i)}$ following [47]:

$$\hat{\boldsymbol{\mu}}_{\mathbf{e}|\boldsymbol{\theta}} = \hat{\boldsymbol{\mu}}_{\mathbf{e}} + \hat{\boldsymbol{\Gamma}}_{\mathbf{e}\boldsymbol{\theta}}\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\theta}}^{-1}(\boldsymbol{\theta} - \hat{\boldsymbol{\mu}}_{\boldsymbol{\theta}}), \tag{3.7}$$

$$\hat{\boldsymbol{\Gamma}}_{\mathbf{e}|\boldsymbol{\theta}} = \hat{\boldsymbol{\Gamma}}_{\mathbf{e}} - \hat{\boldsymbol{\Gamma}}_{\mathbf{e}\boldsymbol{\theta}}\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\theta}}^{-1}\hat{\boldsymbol{\Gamma}}_{\mathbf{e}\boldsymbol{\theta}}^T, \tag{3.8}$$

where

$$\hat{\boldsymbol{\mu}}_{\mathbf{e}} = \frac{1}{n_t}\sum_{j}^{n_t}\mathbf{e}^{(j)}, \tag{3.9}$$

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{\theta}} = \frac{1}{n_t}\sum_{j}^{n_t}\boldsymbol{\theta}_t^{(j)}, \tag{3.10}$$

$$\hat{\boldsymbol{\Gamma}}_{\mathbf{e}} = \frac{1}{n_t - 1}\sum_{j}^{n_t}(\mathbf{e}^{(j)} - \hat{\boldsymbol{\mu}}_{\mathbf{e}})(\mathbf{e}^{(j)} - \hat{\boldsymbol{\mu}}_{\mathbf{e}})^T + \boldsymbol{\Sigma}, \tag{3.11}$$

$$\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\theta}} = \frac{1}{n_t - 1}\sum_{j}^{n_t}(\boldsymbol{\theta}_t^{(j)} - \hat{\boldsymbol{\mu}}_{\boldsymbol{\theta}})(\boldsymbol{\theta}_t^{(j)} - \hat{\boldsymbol{\mu}}_{\boldsymbol{\theta}})^T, \tag{3.12}$$

$$\hat{\boldsymbol{\Gamma}}_{\mathbf{e}\boldsymbol{\theta}} = \frac{1}{n_t - 1}\sum_{j}^{n_t}(\mathbf{e}^{(j)} - \hat{\boldsymbol{\mu}}_{\mathbf{e}})(\boldsymbol{\theta}_t^{(j)} - \hat{\boldsymbol{\mu}}_{\boldsymbol{\theta}})^T. \tag{3.13}$$

The diagonal matrix $\boldsymbol{\Sigma}$ corresponds to an additional white noise component that is added to $\hat{\boldsymbol{\Gamma}}_{\mathbf{e}}$ to improve computational stability and to compensate for underestimation of error variance due to a limited training set size. The standard deviations in $\boldsymbol{\Sigma}$ (i.e. square roots of the diagonal elements) are chosen to be 10% of the sample mean of $\mathbf{Z}_t$.

Using $\hat{\mathbf{A}}$, $\boldsymbol{\mu}_{\mathbf{e}|\boldsymbol{\theta}}$, and $\boldsymbol{\Gamma}_{\mathbf{e}|\boldsymbol{\theta}}$, the following Gaussian likelihood function is formulated:

$$\pi(\mathbf{z}|\boldsymbol{\theta}, \hat{\mathbf{A}}) = \mathcal{N}(\mathbf{z}|\hat{\mathbf{A}}\boldsymbol{\phi}(\boldsymbol{\theta}) + \hat{\boldsymbol{\mu}}_{\mathbf{e}|\boldsymbol{\theta}}, \hat{\boldsymbol{\Gamma}}_{\mathbf{e}|\boldsymbol{\theta}}), \tag{3.14}$$

where $\hat{\mathbf{A}}\boldsymbol{\phi}(\boldsymbol{\theta}) + \hat{\boldsymbol{\mu}}_{\mathbf{e}|\boldsymbol{\theta}}$ is the expectation and $\hat{\boldsymbol{\Gamma}}_{\mathbf{e}|\boldsymbol{\theta}}$ is the covariance. Notation $\mathcal{N}(\,\cdot\,)$ refers here to the probability density function of the multivariate normal distribution.

In the publication **I**, only a positivity constraint is used as a prior. Thus, the posterior density is

$$\pi(\boldsymbol{\theta}|\mathbf{z}) \propto \begin{cases} \mathcal{N}(\mathbf{z}|\hat{\mathbf{A}}\boldsymbol{\phi}(\boldsymbol{\theta}) + \hat{\boldsymbol{\mu}}_{\mathbf{e}|\boldsymbol{\theta}}, \hat{\boldsymbol{\Gamma}}_{\mathbf{e}|\boldsymbol{\theta}}) & \boldsymbol{\theta} \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3.15}$$

The stand attributes, such as tree height and stem volume, are strongly correlated. Thus, in the extended studies where these attributes were included in $\boldsymbol{\theta}$, the correlations needed to be modeled in the prior distribution. Thus, in addition to the positivity constraint, a Gaussian prior is introduced. The prior mean and covariance are learned from the training data and the prior density is:

$$\pi(\boldsymbol{\theta}) \propto \begin{cases} \mathcal{N}(\boldsymbol{\theta}|\hat{\boldsymbol{\mu}}_{\boldsymbol{\theta}}, \hat{\boldsymbol{\Gamma}}_{\boldsymbol{\theta}}) & \boldsymbol{\theta} \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3.16}$$

The posterior density is then

$$\pi(\boldsymbol{\theta}|\mathbf{z}) \propto \begin{cases} \mathcal{N}(\mathbf{z}|\hat{\mathbf{A}}\boldsymbol{\phi}(\boldsymbol{\theta}) + \hat{\boldsymbol{\mu}}_{\mathbf{e}|\boldsymbol{\theta}}, \hat{\boldsymbol{\Gamma}}_{\mathbf{e}|\boldsymbol{\theta}})\mathcal{N}(\boldsymbol{\theta}|\hat{\boldsymbol{\mu}}_{\boldsymbol{\theta}}, \hat{\boldsymbol{\Gamma}}_{\boldsymbol{\theta}}) & \boldsymbol{\theta} \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3.17}$$

3.2.1 Dominant tree species identification

In the publication **I**, the posterior density of stem volume vector $\mathbf{v}$ was also used to calculate a class uncertainty measure for dominant tree species identification. From the posterior density $\pi(\mathbf{v}|\mathbf{z}, \hat{\mathbf{A}})$, the (posterior) probability that the i'th species is dominant ($p_{\text{dominance},i}$) is:

$$\begin{aligned} p_{\text{dominance},i} &= \int_G \pi(\mathbf{v}|\mathbf{z}, \hat{\mathbf{A}})\mathrm{d}\mathbf{v}, \\ &\text{where } G = \{\mathbf{v} \in \mathbb{R}^3 | \mathbf{v}_i > \mathbf{v}_j,\ \forall i \neq j\}. \end{aligned} \tag{3.18}$$

A high dominance probability $p_{\text{dominance},i}$ (for example over 0.95), implies that the species i is very likely the correct dominant species, while a low value of the largest $p_{\text{dominance},i}$ (for example ca. 0.5) indicates a high uncertainty in the identified dominant species. Based on this idea, $p_{\text{dominance}}$ was used to trace plots that have a high risk of misidentifying the dominant species. The dominance probability was compared with the heuristic of plot-level mixedness of a point estimate.

3.2.2 Computation, estimates, and the reference method

Because of the nonlinearity caused by the inclusion of relative stem volumes in $\boldsymbol{\phi}(\boldsymbol{\theta})$ and the positivity constraint, closed form expressions for the posterior mean, posterior covariance or credible intervals are not available. The mean and credible intervals were thus solved numerically using the DRAM (delayed rejection adaptive Metropolis) algorithm. In the results of the publication **I**, 120000 Monte Carlo samples per plot were computed, using 12 parallel chains of 10000 samples each (the number does not include the 1000 sample burn-in period). For the values of other DRAM

algorithm parameters, see the publication **I**. For the extended results, 100000 sample chains (1.2 million samples) were used due to the increased number of variables.

The DRAM samples were used to compute 95% credible intervals using the narrowest interval criterion, where the optimization problem was solved using a brute force algorithm. The DRAM sample with the highest posterior probability density value was used as an approximation to the posterior mode (MAP estimate), which was used as the point estimate. The credible intervals for the total N, BA and V were computed by first summing up the species-specific DRAM samples to form samples of the total variables and then computing the CIs. The point estimates for the total variables are simply sums of the species-specific MAP estimates.

As a reference method, a kNN-based approach is used. We select ten predictors from the data using a simulated annealing -based optimization approach presented in [55] and use the most similar neighbor (MSN) method for the neighbor selection. The number of neighbors is chosen to be 5, as in [55]. The predictor selection is done using the whole data set and leave-one-out (LOO) cross-validation.

3.3 REVIEW OF THE RESULTS

Figure 3.1 shows an example of posterior marginal densities (histograms of the Monte Carlo samples) for the species-specific stem volumes and total stem volume of a single test plot from the publication **I**. The point estimate values and the field measured values for the variables as well as the 95% credible intervals (shaded areas) are illustrated in the figure. In this example, the posterior marginals for pine and spruce are widened, while the posterior marginal of deciduous trees is relatively sharp. The wider posterior marginals imply higher uncertainty in the pine and spruce volumes than in the deciduous volume.

Figure 3.2 shows the posterior marginal densities for species-specific and total stem volumes for the same example plot as in Figure 3.1 from the extended results, where the additional attributes were included in $\boldsymbol{\theta}$. In general, the posterior marginals are similar to the article **I** results: marginals for pine and spruce are wide and deciduous marginal is narrow. Two significant differences between the figures can be, however, observed. First, the marginals for pine and spruce are not as wide in the extended results as in Figure 3.1. Secondly, the marginal of deciduous volume does not tend to zero as strongly. The first difference is most probably the result of better data fit from the higher dimensional model, that is, the predictor variation can be more accurately explained when using multiple stand attributes than with only the stem volume. The second difference is most likely caused by the introduction of a Gaussian prior.

Figure 3.3 shows joint marginal densities of stem volume for each species pair in the same example plot from the publication **I**. These images illustrate the joint posterior statistics of the species pairs; for example, the elongated shape of the posterior marginal in the pine-spruce plane reveals the uncertainty in distinguishing between pine and spruce. Figure 3.4 contains the corresponding 2D marginals from the extended results. The elongation direction of the 2D marginals stays the same, yet the uncertainty is diminished.

Figure 3.5 shows 2D marginals between pine height, pine stem diameter, and pine stem number, for a different example plot. The joint marginal density of height and diameter is strongly correlated. The stem number on the other hand shows no large posterior correlation with either height or diameter, while such correlation was

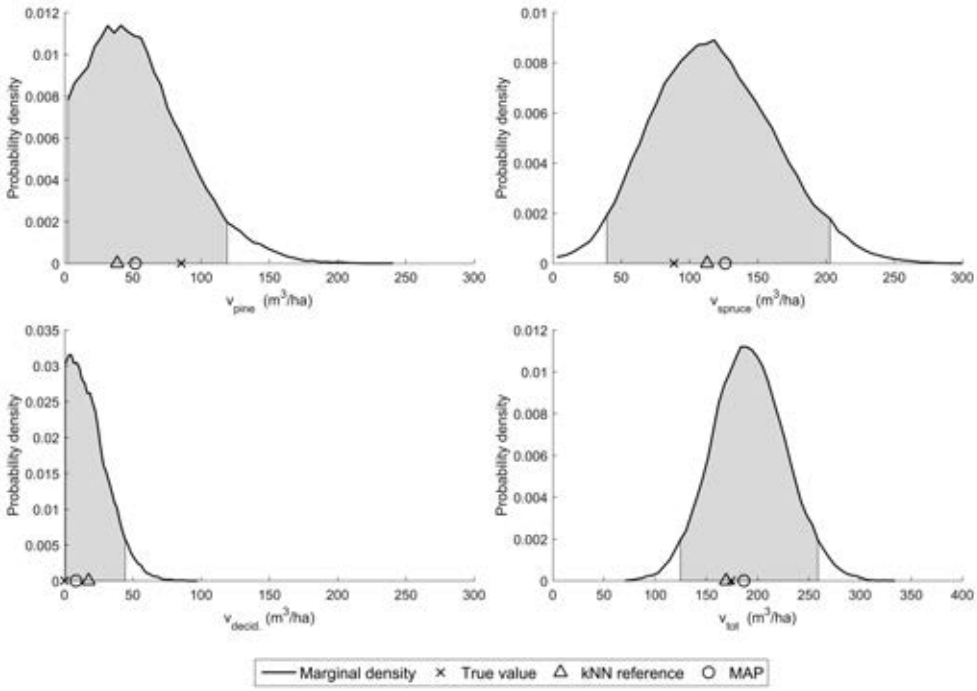

Figure 3.1: Examples of marginal densities of species-specific and total stem volume of a single test plot from the publication **I**. The 95% credible intervals are shaded with gray. Top left: pine, top right: spruce, bottom left: decidous trees, bottom right: total volume. © 2016 IEEE.

expected to show based on the prior. The elongation in N_{pine} direction implies large uncertainty in stem number estimation.

Tables 3.1–3.3 contain, respectively, the aggregated RMSE% (relative root mean square error), bias% (relative bias), and CI% (credible interval coverage) results from the article **I** and the extended results. The CI coverage is defined as the percentage of plots on which the field-measured value is within the computed 95% credible interval, and is thus a measure of the uncertainty quantification performance. While Bayesian credible intervals do not have the same frequency interpretation as frequentist confidence intervals, if the statistical modeling is correct, then, for example, 95% CI% should tend to 95% [51]. The cross-validation method used to compute the results was leave-one-out.

The Bayesian species-specific stem volume estimates of the publication **I** are somewhat worse in RMSE (Table 3.1) than the corresponding kNN predictions. In total stem volume, the Bayesian method produced better results. The same behavior occurs in the extended results. However, in the Bayesian point estimates, inclusion of more variables improves the RMSE accuracy of the species-specific stem volumes, especially $v_{\text{decid.}}$. For kNN, the performance stays nearly the same when more variables are added. For the other stand attributes, the scenario is closely similar: kNN has better RMSE performance in general when more variables are added, yet for minority species (spruce and deciduous) and total values, the Bayesian method is often nearly equal to or slightly better than kNN (total basal area and total volume).

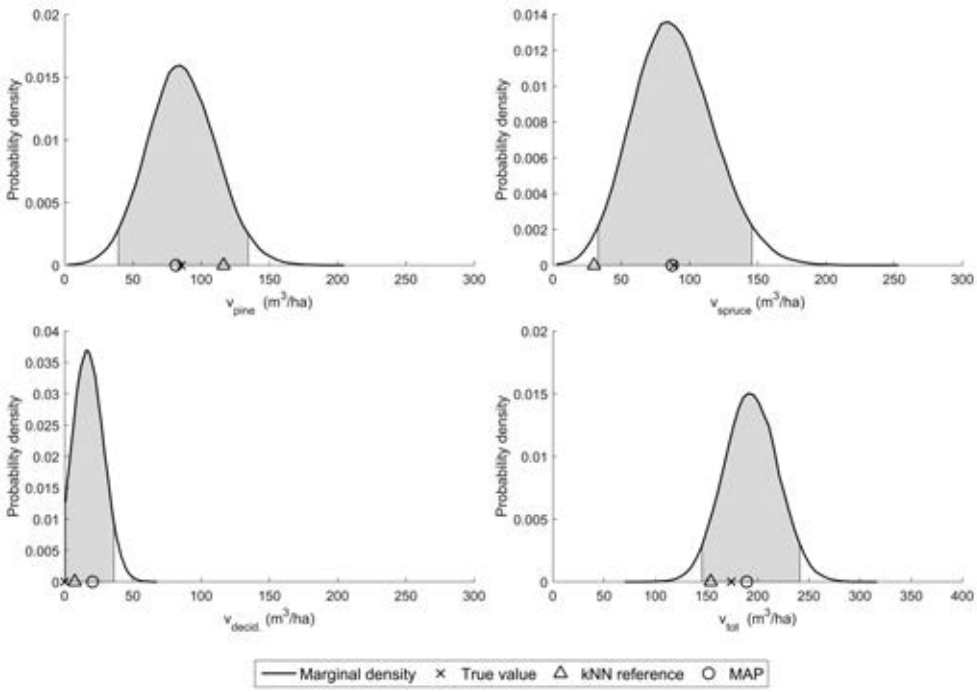

Figure 3.2: Examples of marginal densities of species-specific and total stem volume of the single test plot of Figure 3.1 from the extended results. The 95% credible intervals are shaded with gray. Top left: pine, top right: spruce, bottom left: decidous trees, bottom right: total volume.

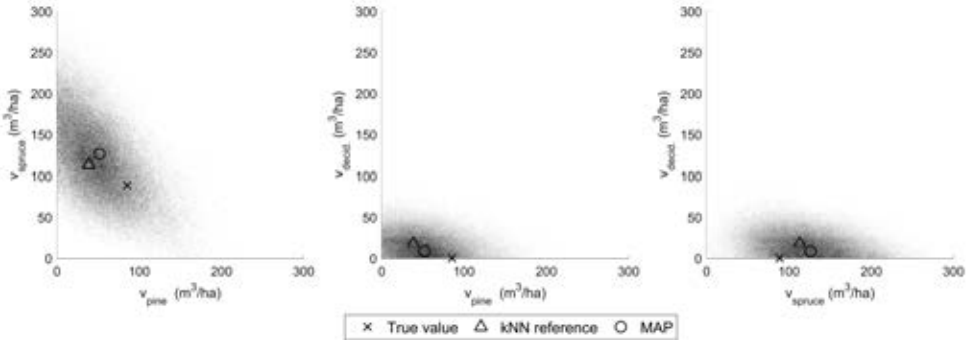

Figure 3.3: Example of joint marginal densities of species-specific volumes for each species pair of a single test plot computed as 2D histograms of MCMC samples from the publication **I**. Darker gray means higher relative probability density. © 2016 IEEE.

Table 3.1: Relative RMSE of the Bayesian and reference kNN estimates. Cross-validation using LOO. The table contains both the results from the publication **I** and extended results with more estimated variables.

	Bayes		kNN	
	Article **I**	Extended	Article **I**	Extended
RMSE%				
h_{pine}		41.28		36.57
h_{spruce}		57.60		54.21
$h_{\text{decid.}}$		72.75		70.43
d_{pine}		36.34		32.99
d_{spruce}		53.28		51.36
$d_{\text{decid.}}$		60.86		59.75
N_{pine}		49.05		41.92
N_{spruce}		84.43		63.78
$N_{\text{decid.}}$		98.67		87.03
N_{total}		33.61		30.15
BA_{pine}		36.31		31.83
BA_{spruce}		70.07		61.66
$BA_{\text{decid.}}$		82.45		75.56
BA_{total}		18.76		20.49
v_{pine}	42.41	39.98	33.30	36.03
v_{spruce}	82.21	80.89	73.72	71.24
$v_{\text{decid.}}$	97.77	88.95	83.69	83.37
v_{total}	21.30	22.09	23.55	24.51

In the relative bias (Table 3.2) results, kNN is again superior. The Bayesian estimates tend to underestimate the volume of the most common tree species (Scots pine) and overestimate the volume of the minority species, following a similar error structure that has been noticed in other studies on forest inventory [56]. It is evident from Table 3.2, that similar error structure occurs also in the other stand attributes.

In CI coverage, the results from the article **I** were excellent, with CI% close to the ideal 95%. In the extended results, CI% for the stem volumes drops to a bit over 80%, see Table 3.3. The CI% for the additional stand attributes are roughly in the same range. Two possible reasons for this reduction in CI%, while RMSE and bias stay roughly the same, are multicollinearity, i.e. the strong correlation between different stand attributes, and the so-called curse of dimensionality, that is, when the dimension increases, the volume spanned by the training set is relatively smaller and smaller compared to the space of possible solutions.

In the publication **I**, the effect of training set cover was studied by computing error metrics separately for the plots that were inside the convex hull of the training set and for those outside. It was found that the performance in all metrics was significantly better for the plots that were inside the convex hull. Publication **I** also includes results for multiple training set sizes; these are summarized in Figure 3.6.

Table 3.2: Relative bias of the Bayesian and reference kNN estimates. Cross-validation using LOO. The table contains both the results from the publication **I** and extended results with more estimated variables.

	Bayes		kNN	
	Article **I**	Extended	Article **I**	Extended
bias%				
h_{pine}		-0.22		-0.09
h_{spruce}		2.25		1.51
$h_{\text{decid.}}$		3.08		-0.03
d_{pine}		0.16		0.60
d_{spruce}		2.93		1.20
$d_{\text{decid.}}$		2.83		0.30
N_{pine}		-3.65		1.86
N_{spruce}		13.06		1.38
$N_{\text{decid.}}$		13.45		-0.82
N_{total}		5.39		1.12
BA_{pine}		-3.94		2.43
BA_{spruce}		7.33		1.94
$BA_{\text{decid.}}$		11.69		-5.20
BA_{total}		1.16		1.36
v_{pine}	-7.59	-3.38	0.57	2.01
v_{spruce}	9.04	4.17	-2.57	1.42
$v_{\text{decid.}}$	14.00	10.70	-0.26	-7.20
v_{total}	-0.59	0.24	-0.40	0.87

In summary, the RMSE increases and CI% grows smaller when the training set size decreases.

The dominant tree species identification (see Section 3.2.1) that was done in the publication **I** showed that the Bayesian approach was slightly better at identifying the plots with the highest risk of misidentification than the kNN based heuristic. The classification uncertainty method could also be used for example on thinning need assessment, using estimated stem number and dominant height.

3.4 DISCUSSION

The poorer RMSE performance of the Bayesian approach stems most likely from the use of a nearly linear model. While the model has attractive conservative extrapolation properties, it fails to model the nonlinear relationship between the predictors and the stand attributes inside the training set. Thus, in the future, it would be useful to consider more complicated models, even if increased model complexity often introduces additional problems with overfitting. In addition to more complex models, more advanced methods of fitting the model to the training data might improve the results; in the present work, the model was fitted using ordinary least squares.

Table 3.3: CI coverage of the Bayesian estimates. Cross-validation using LOO. The table contains both the results from the publication **I** and extended results with more estimated variables.

	Bayes	
	Article **I**	Extended
CI%		
h_{pine}		86.61
h_{spruce}		84.99
$h_{\text{decid.}}$		83.37
d_{pine}		86.21
d_{spruce}		77.48
$d_{\text{decid.}}$		76.47
N_{pine}		87.01
N_{spruce}		85.60
$N_{\text{decid.}}$		91.08
N_{total}		78.09
BA_{pine}		77.08
BA_{spruce}		72.21
$BA_{\text{decid.}}$		78.09
BA_{total}		85.80
v_{pine}	93.51	81.74
v_{spruce}	96.15	83.98
$v_{\text{decid.}}$	97.36	83.16
v_{total}	94.93	83.40

The extended results presented in this chapter were well in line with the published results of the article **I**, and answer some of the open questions presented in the discussion of the publication **I**. The number of estimated stand attributes can be increased, which results in slightly better estimation accuracy, but poorer uncertainty quantification performance. Using a Gaussian prior density that is constructed using the training data is a feasible approach and the prior does not introduce large biases in the estimates.

Perhaps the most valuable property of the approach described here is the great flexibility of the Bayesian framework. The framework offers a consistent and mathematically justified way to include auxiliary information in the estimation of stand attributes. These auxiliary data could be for example measurements from previous campaigns on the same location (see e.g. [57]) or results from other remote sensing modalities.

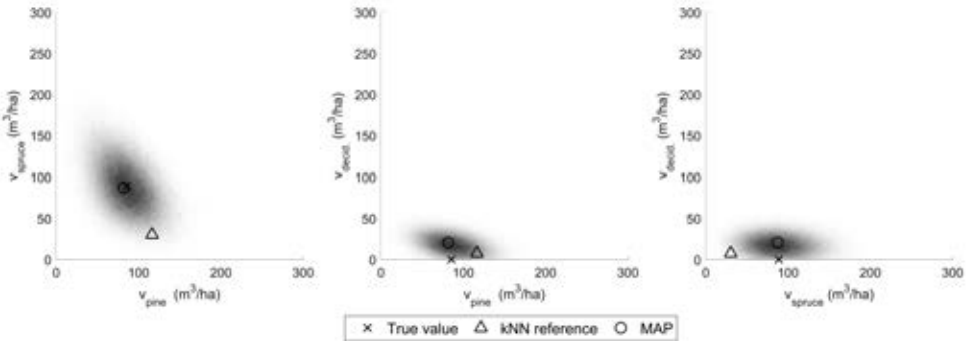

Figure 3.4: Example of joint marginal densities of species-specific volumes for each species pair of a single test plot computed as 2D histograms of MCMC samples from the extended results. Darker gray means higher relative probability density.

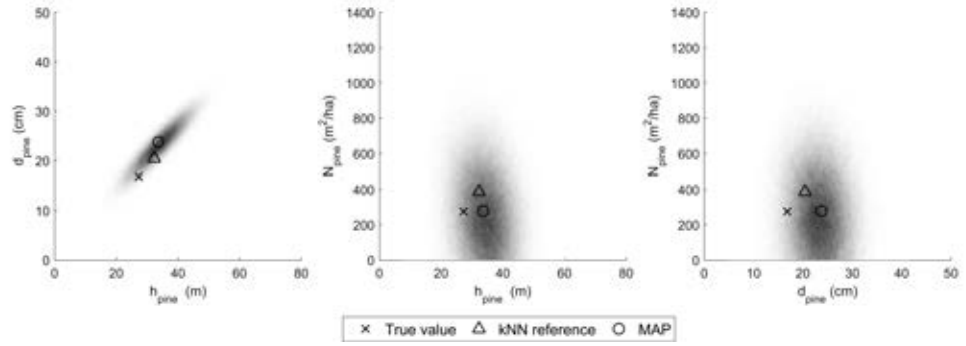

Figure 3.5: Example of joint marginal densities of pine height, pine stem diameter, and pine stem number of a single test plot computed as 2D histograms of MCMC samples from the extended results. Darker gray means higher relative probability density.

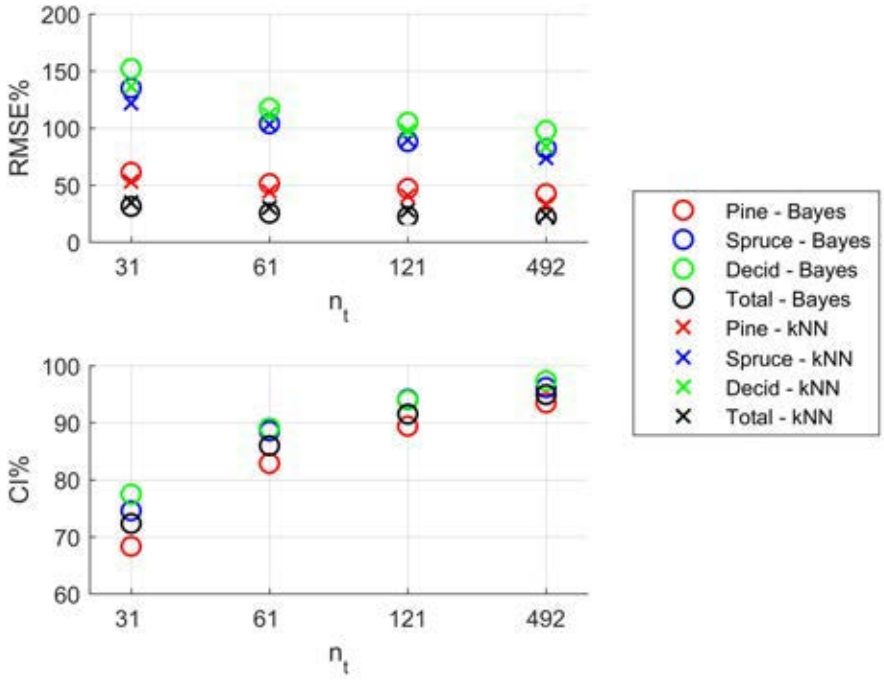

Figure 3.6: Evolution of RMSE% and CI% of species-specific stem volumes with training set size n_t (Table 2 of the publication **I**).

4 Bayesian estimation of canopy leaf area index from satellite measurements

In this chapter, the Bayesian forest reflectance model inversion and results of the publications **II** and **III** are reviewed and summarized. Publication **II** introduced the methodology, which was then evaluated using simulated forest reflectance data. In the publication **III** the method was applied to real EO-1 Hyperion data.

4.1 FOREST REFLECTANCE MODEL

In this thesis, forest reflectance spectra are modeled using the PARAS forest reflectance model [58] which is based on the concept of photon recollision probability. The PARAS model was chosen because it has the advantage of containing relatively few independent variables and performing well in boreal forests [58]. The bidirectional reflectance factor (BRF) of a forest, $r(\theta_1, \theta_2, \lambda)$, for a given solar zenith angle θ_1, viewing zenith angle θ_2, and wavelength λ, is modeled as:

$$r(\theta_1, \theta_2, \lambda) = \rho_g(\theta_1, \theta_2, \lambda) t_c(\theta_1) t_c(\theta_2) + f(\theta_1, \theta_2, \lambda) i_c(\theta_1) \frac{\omega_L(\lambda) - p\omega_L(\lambda)}{1 - p\omega_L(\lambda)}, \quad (4.1)$$

where ρ_g is the BRF of the understory layer, $t_c(\theta)$ is the tree canopy transmittance in the direction θ, $i_c(\theta) = 1 - t_c(\theta)$ is the canopy interceptance, $f(\theta_1, \theta_2, \lambda)$ the canopy upward scattering phase function, and $\omega_L(\lambda)$ the leaf single scattering albedo. The photon recollision probability p is used in the model to describe the aggregated structure of the forest canopy [42,59]. It is the probability that a photon, after having survived an interaction with a canopy element, will interact with the canopy again.

Effective leaf area index (LAI) is used in the reflectance model in order to make the estimated quantities comparable with the optical LAI field measurements. The effective LAI is assumed to follow the model $\mathrm{LAI}_{\mathrm{eff}} = \beta \mathrm{LAI}$, where β is the shoot clumping factor. The photon recollision probability p is approximated following [60] as

$$p = 1 - \frac{1 - t_d}{\mathrm{LAI}} = 1 - \frac{\beta(1 - t_d)}{\mathrm{LAI}_{\mathrm{eff}}}, \quad (4.2)$$

where t_d is the diffuse transmittance for the canopy layer. The canopy transmittance is modeled using Beer-Lambert's law as

$$t_c(\theta) = \exp\left(-\frac{\beta}{2} \frac{\mathrm{LAI}_{\mathrm{eff}}}{\cos(\theta)}\right), \quad (4.3)$$

from which the diffuse canopy transmittance t_d in the equation (4.2) is calculated following [61]:

$$t_d = 2 \int_0^{\frac{\pi}{2}} t_c(\theta) \cos(\theta) \sin(\theta) \mathrm{d}\theta. \quad (4.4)$$

The upward scattering phase function $f(\theta_1, \theta_2, \lambda)$ is approximated using the proportion of upward scattered radiation Q as [62]

$$f(\theta_1, \theta_2, \lambda) \approx Q = \frac{1}{2} + \frac{q}{2} \frac{1 - p\omega_L}{1 - pq\omega_L}, \tag{4.5}$$

where q in is a wavelength independent semi-empirical scattering asymmetry parameter. Parameter q describes the decrease in probability of the photon escaping the canopy with increasing scattering order, in other words, it models how photon escape probability decreases as the photon scatters deeper inside the canopy. Thus q is related to canopy density and increases with LAI (Table 2 in [62]).

4.1.1 Wavelength dependence

Leaf albedo ω_L and understory reflectance ρ_g are wavelength dependent parameters. Thus, in the model, ω_L and ρ_g are vectors of the same length as the satellite-measured data vector. To reduce the number of unknown variables in the inverse problem, we utilize known features of the vegetation spectra: The (green) vegetation spectra have a typical shape and stucture, which enables the use of reduced order parametric representations for ω_L and ρ_g. In the publications **II** and **III**, cubic monotone Hermite splines are used to represent the spectral variables using 27 manually chosen node points (see Figure 1 of the publication **II**). The node points are chosen in a way that allows the spline representation to follow the typical shape of a vegetation spectrum. The approximation is generally accurate (see Figure 1 in the publication **II**), but inevitably introduces some error.

Using the spline, the variables ω_L and ρ_g are rewritten as

$$\omega_L = S(\lambda; \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\omega}}_L), \tag{4.6}$$

$$\rho_g = S(\lambda; \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\rho}}_g), \tag{4.7}$$

where $S(\,\cdot\,)$ is the spline function (piecewise polynomial), $\tilde{\boldsymbol{\lambda}} \in \mathbb{R}^{27}$ is a vector consisting of wavelengths corresponding to the spline nodes, and $\tilde{\boldsymbol{\omega}}_L \in \mathbb{R}^{27}$ and $\tilde{\boldsymbol{\rho}}_g \in \mathbb{R}^{27}$, respectively, are the values of ω_L and ρ_g at the node points $\tilde{\boldsymbol{\lambda}}$. Because $\tilde{\boldsymbol{\lambda}}$ is fixed, the spline approximations (4.6) and (4.7) are fully determined by $\tilde{\boldsymbol{\omega}}_L$ and $\tilde{\boldsymbol{\rho}}_g$, respectively. Thus, using the spline approximations, the low-dimensional vectors $\tilde{\boldsymbol{\omega}}_L$ and $\tilde{\boldsymbol{\rho}}_g$ are substituted for full-length $\boldsymbol{\omega}_L$ and $\boldsymbol{\rho}_g$ as variables in the reflectance model.

4.2 BAYESIAN FORMULATION

The measurement vector $\mathbf{z} \in \mathbb{R}^{150}$ is here a vector consisting of the measured BRFs on all wavelength bands. The vector of unknown variables is $\boldsymbol{\theta} = \begin{bmatrix} \mathrm{LAI}_{\mathrm{eff}} & \tilde{\boldsymbol{\omega}}_L^T & \tilde{\boldsymbol{\rho}}_g^T & \beta \end{bmatrix}^T \in \mathbb{R}^{56}$. The measurement $\mathbf{z}$ is modeled as

$$\mathbf{z} = h(\boldsymbol{\theta}) + \mathbf{e}, \tag{4.8}$$

where $h(\boldsymbol{\theta})$ is the PARAS model (4.1), including the substituted approximations for ω_L, ρ_g,t_c, Q, and p, and e is an additive Gaussian error term. With this model, the

likelihood function $\pi(\mathbf{z}|\boldsymbol{\theta})$ is of the form

$$\pi(\mathbf{z}|\boldsymbol{\theta}) \propto \exp\left(-\frac{1}{2}(\mathbf{z}-h(\boldsymbol{\theta}))^T\boldsymbol{\Gamma}_\mathbf{e}^{-1}(\mathbf{z}-h(\boldsymbol{\theta}))\right), \tag{4.9}$$

where $\boldsymbol{\Gamma}_\mathbf{e}$ is the covariance matrix of $\mathbf{e}$.

In the publication **II**, $\mathbf{e}$ is assumed to have zero mean and a standard deviation of 10% of the measurement $\mathbf{z}$ (i.e. the same error variance that was used to generate the simulated measurements). In the publication **III**, $\mathbf{e}$ is assumed to have zero mean and a standard deviation of 5% of the average reflectance $\mathbf{z}$ in wavelengths 488–691 nm, 702–1346 nm, and 1477–1800 nm, and 10% of the average reflectance in wavelengths 2032–2355 nm. The standard deviation values were chosen based on the EO-1 Hyperion radiometric accuracy [63] with some extra deviation to compensate for possible uncertainty resulting from the preprocessing and atmospheric correction.

4.2.1 Prior density

Uniform prior distributions are set for $\mathrm{LAI}_{\mathrm{eff}}$ and β: $\mathrm{LAI}_{\mathrm{eff}}$ is by definition non-negative and unrealistically large values ($\mathrm{LAI}_{\mathrm{eff}} > 10$) are constrained out

$$\pi(\mathrm{LAI}_{\mathrm{eff}}) = \begin{cases} \frac{1}{10}, & 0 \le \mathrm{LAI}_{\mathrm{eff}} \le 10 \\ 0, & \text{otherwise,} \end{cases} \tag{4.10}$$

and β is constrained to the empirically observed range $[0.4, 1]$ [64]

$$\pi(\beta) = \begin{cases} \frac{5}{3}, & 0.4 \le \beta \le 1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.11}$$

The spectral variables $\tilde{\boldsymbol{\omega}}_L$ and $\tilde{\boldsymbol{\rho}}_g$ are modeled by truncated multivariate Gaussian prior distributions:

$$\pi(\tilde{\boldsymbol{\omega}}_L) \propto \begin{cases} \exp\left(-\frac{1}{2}(\tilde{\boldsymbol{\omega}}_L - \boldsymbol{\mu}_{\tilde{\omega}_L})^T\boldsymbol{\Gamma}_{\tilde{\omega}_L}^{-1}(\tilde{\boldsymbol{\omega}}_L - \boldsymbol{\mu}_{\tilde{\omega}_L})\right), & 0 \le \tilde{\boldsymbol{\omega}}_L \le 1 \\ 0, & \text{otherwise,} \end{cases} \tag{4.12}$$

$$\pi(\tilde{\boldsymbol{\rho}}_g) \propto \begin{cases} \exp\left(-\frac{1}{2}(\tilde{\boldsymbol{\rho}}_g - \boldsymbol{\mu}_{\tilde{\rho}_g})^T\Gamma_{\tilde{\boldsymbol{\rho}}_g}^{-1}(\tilde{\boldsymbol{\rho}}_g - \boldsymbol{\mu}_{\tilde{\rho}_g})\right), & 0 \le \tilde{\boldsymbol{\rho}}_g \le 1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.13}$$

The expected values $\boldsymbol{\mu}_{\tilde{\omega}_L}$ and $\boldsymbol{\mu}_{\tilde{\rho}_g}$ are derived from published measurement data: the leaf albedo data from [65] and the understory reflectance data from [66]. The covariance matrices $\boldsymbol{\Gamma}_{\tilde{\omega}_L}$ and $\boldsymbol{\Gamma}_{\tilde{\rho}_g}$ are contructed by setting the standard deviation to a certain fraction (20% in **II**, 25% in **III**) of the prior expectation on each channel and using a preconstructed correlation matrix. See the publication **II** for details.

The possible correlations between the variables $\mathrm{LAI}_{\mathrm{eff}}$, β, $\tilde{\boldsymbol{\omega}}_L$, and $\tilde{\boldsymbol{\rho}}_g$ are ignored, because quantified information on these correlations is not available. Therefore it is approximated that these variables are mutually independent and the resulting prior density for $\boldsymbol{\theta}$ is

$$\pi(\boldsymbol{\theta}) = \pi(\mathrm{LAI}_{\mathrm{eff}})\pi(\tilde{\boldsymbol{\omega}}_L)\pi(\tilde{\boldsymbol{\rho}}_g)\pi(\beta). \tag{4.14}$$

The posterior density $\pi(\boldsymbol{\theta}|\mathbf{z})$ is finally obtained by combining the likelihood (4.9) and the prior (4.14) using the Bayes' theorem (2.5).

4.2.2 Bayesian estimates

From the posterior density, the posterior conditional mean estimates and 95% credible intervals were computed using DRAM [54]. In the publication **II**, total of 600000 Monte Carlo samples were computed for each stand, using 12 parallel sampling chains of 50000 samples each. In the publication **III**, this was increased to 840000 samples (12 parallel chains of 70000 samples). The numbers do not include the 5000 burn-in period in the beginning of each chain. For a more detailed description, see the publication **II**.

The narrowest interval 95% CIs were used to evaluate the CI performance (as in the publication **I**). In addition, equal tails CIs were computed to visualize the posterior marginal densities of ρ_g and ω_L.

In the publication **II**, Bayesian estimates using a uniform prior were also computed to evaluate the effect of the prior model on the results. The uniform prior has the same box constraint as in the prior (4.14), that is, $\mathrm{LAI}_{\mathrm{eff}}$ is constrained to $[0, 10]$, β to $[0.4, 1]$, and ρ_g and ω_L to $[0, 1]$.

4.3 MATERIALS

4.3.1 Simulation study

In the simulation study **II**, a total of 500 random simulated boreal forest stands were generated. First, the dominant tree species (pine, spruce or broadleaved) was chosen. The proportion of the dominant species in the species mixture was sampled uniformly from the interval 50%–100%; the remainder was then randomly divided between the two minority species. The composition of the understory layer was then sampled to roughly emulate the typical species composition of a Finnish boreal forest with the chosen dominant tree species. Leaf area index was sampled from a uniform density, β and ω_L were generated based on the sampled tree species composition, and ρ_g based on the sampled understory composition. See the publication **II** for details.

After the vector of input parameters was sampled, the PARAS model was used to generate the simulated BRF. Gaussian random noise with standard deviation of 10% of the reflectance on each band was added to the modelled reflectance.

4.3.2 Hyytiälä data set

In the publication **III**, a set of real data was used. The study area is located next to Hyytiälä Forestry Field Station in southern Finland. The data consist of 18 stands with different species compositions and age classes typical to the region. Field measurements of $\mathrm{LAI}_{\mathrm{eff}}$ were carried out in early May, early June and early July in 2010. Concurrently with the LAI measurements, data on understory reflectance spectra was collected in four stands representing common site fertility types: mesic, xeric, sub-xeric and herb-rich sites. In addition, we had access to regular stand inventory data which had been collected in all our study plots a year before the satellite images were acquired.

Three EO-1 Hyperion satellite images were acquired from our study area concurrently with the field data collection. Hyperion was a narrowband imaging spectrometer onboard NASA's Earth Observing-One (EO-1) with 242 spectral bands (356 – 2577 nm), of which 150 were used in the study, and a 30 m $\times$ 30 m spatial resolution [63].

The set of Hyperion images captures the main phenological changes occurring in the study area in 2010: the image from May corresponds to the time of bud burst, the image from June to the full leaf-out situation, and the image from July to the time of maximal leaf area. The mean HDRFs for each study stand were extracted using a 3×3 pixel window which corresponds to the area covered by the field measurements in each stand. See the publication **III** for a detailed description of the field measurements and data processing.

4.4 REFERENCE METHODS

In the publication **II** the Bayesian estimates were compared with two reference methods: 1) The maximum likelihood (ML) estimates of $\mathrm{LAI}_{\mathrm{eff}}$ when parameters ω_L, ρ_g and β were fixed to their population means, and 2) empirical linear regression with a narrow-band vegetation index (VI).

In the publication **III**, the linear VI regression was used as the reference method. Two regressions were considered: one using the field-measured data as the training data in a leave-one-out configuration, and the second with simulated training data that was constructed to closely emulate the field data.

4.4.1 Maximum likelihood estimate

The conventional approach to model based estimation of $\mathrm{LAI}_{\mathrm{eff}}$ is to invert the reflectance model after fixing the other model parameters (here ω_L, ρ_g and β) to some predefined values. In the publication **II**, conventional maximum likelihood (ML) estimates were computed by maximizing the likelihood function (4.9) with respect to $\mathrm{LAI}_{\mathrm{eff}}$, with ω_L, ρ_g, and β fixed.

The tolerance of such $\mathrm{LAI}_{\mathrm{eff}}$ estimate to misspecification of the parameters ω_L, ρ_g, and β was studied: For each of the 500 simulated study stands, the ML estimate was computed using two choices of parameters ω_L, ρ_g, and β: 1) In the first ML estimate, the true parameter values in the corresponding study stand were used. This choice is of course unrealistic, since these parameters are practically always unknown. 2) In the second set of ML estimates, parameters ω_L, ρ_g and β were fixed to their average values over the ensemble of simulated study stand test, i.e., to their population means. The latter estimate can be considered as a solution corresponding to the best realistically available approximation for the parameter values, and is expected to exhibit estimation error that is caused by the misspecification of the parameters.

The one-dimensional optimization problem (maximizing (4.9) with respect to $\mathrm{LAI}_{\mathrm{eff}}$) was solved by brute force to 0.1% accuracy, to ensure that the resulting estimate was the global maximum (the likelihood function has multiple local maxima in some cases). For computational reasons, the range of $\mathrm{LAI}_{\mathrm{eff}}$ was constrained to $[0, 10]$.

4.4.2 Vegetation index regression

The Bayesian approach was also compared to an empirical linear regression with a narrow-band vegetation index (VI). As there are a wide range of spectral indices applied in hyperspectral remote sensing of vegetation, we selected the simple ratio water index (SRWI) which has recently been reported as the best performing common index for estimating $\mathrm{LAI}_{\mathrm{eff}}$ in our biome of interest, i.e. the boreal forests [67].

Heiskanen *et al.* [67] used the same Hyytiälä data set of processed Hyperion images and ground reference LAI measurements to compare the performance of different spectral vegetation indices in estimating boreal forest LAI. The SRWI is defined as

$$\mathrm{SRWI} = \frac{r_{854}}{r_{1235}}, \tag{4.15}$$

where the subscript refers to the wavelength in nanometers.

In the publication **II**, a separate set of 100 random training stands were generated and the SRWI was calculated for each stand. Ordinary linear regression was then done between $\mathrm{LAI}_{\mathrm{eff}}$ and SRWI in the training set. Finally, the regression model was used to estimate $\mathrm{LAI}_{\mathrm{eff}}$ for the 500 simulated study stands.

In the publication **III**, two reference VI regressions were used: 1) regression where the real measurement data were used as a training set and 2) regression where the training set was simulated using the PARAS model.

In the VI regression based on real training data, SRWI was first calculated for each measured stand. Leave-one-out cross-validation was then done: each study stand was left out at a time and the rest of the data was used as a training set. Ordinary linear regression was done between field-measured $\mathrm{LAI}_{\mathrm{eff}}$ and SRWI in the training set. The regression model was finally used to estimate $\mathrm{LAI}_{\mathrm{eff}}$ for the left out stand.

In the simulation based VI regression, a set of synthetic training data was simulated with the PARAS model using the field-measured $\mathrm{LAI}_{\mathrm{eff}}$, and the known tree species composition and understory type of each stand in the data set. The aim was to construct a simulated training set as close as possible in composition to the data set used in this study. As in the real training data case, ordinary linear regression was done between $\mathrm{LAI}_{\mathrm{eff}}$ and SRWI in the simulated training set, and the regression model was then used to predict $\mathrm{LAI}_{\mathrm{eff}}$ for each stand.

4.5 REVIEW OF THE RESULTS

4.5.1 Simulation study

Posterior marginal density of $\mathrm{LAI}_{\mathrm{eff}}$ and posterior credible intervals for ω_L and ρ_g of two simulated example forest stands are illustrated in Figure 4.1. The first example stand has a low LAI (and thus, a low canopy cover) and the second example has a high LAI and dense canopy. For details, see the publication **II**. These two simulated example stands illustrate three important tendencies in the shape of the posterior density that depend on the value of LAI: 1) The posterior marginal density of $\mathrm{LAI}_{\mathrm{eff}}$ widens as $\mathrm{LAI}_{\mathrm{eff}}$ increases. This results from the saturation of canopy reflectance in dense canopies, which increases uncertainty in the estimate $\mathrm{LAI}_{\mathrm{eff}}$ and thus the posterior variance. 2) The posterior variance of leaf albedo ω_L decreases as $\mathrm{LAI}_{\mathrm{eff}}$ increases, because in a dense canopy most of the reflected light comes from the canopy. 3) The posterior variance of understory reflectance ρ_g increases with increasing $\mathrm{LAI}_{\mathrm{eff}}$ for the same reason; in a dense canopy a minuscule part of the forest reflectance comes from the understory and therefore the reflectance measurement contains little to no information on the understory reflectance.

Scatter plots of the simulation results for the Bayesian approach, the Bayesian approach with uniform prior, the VI regression, the ML estimate using true values for parameters other than LAI, and the ML estimate using mean values are shown in Figure 4.2. Pine, spruce, and deciduous dominated stands are shown with different symbols. Starting with the ML estimate using the true parameter values (bottom

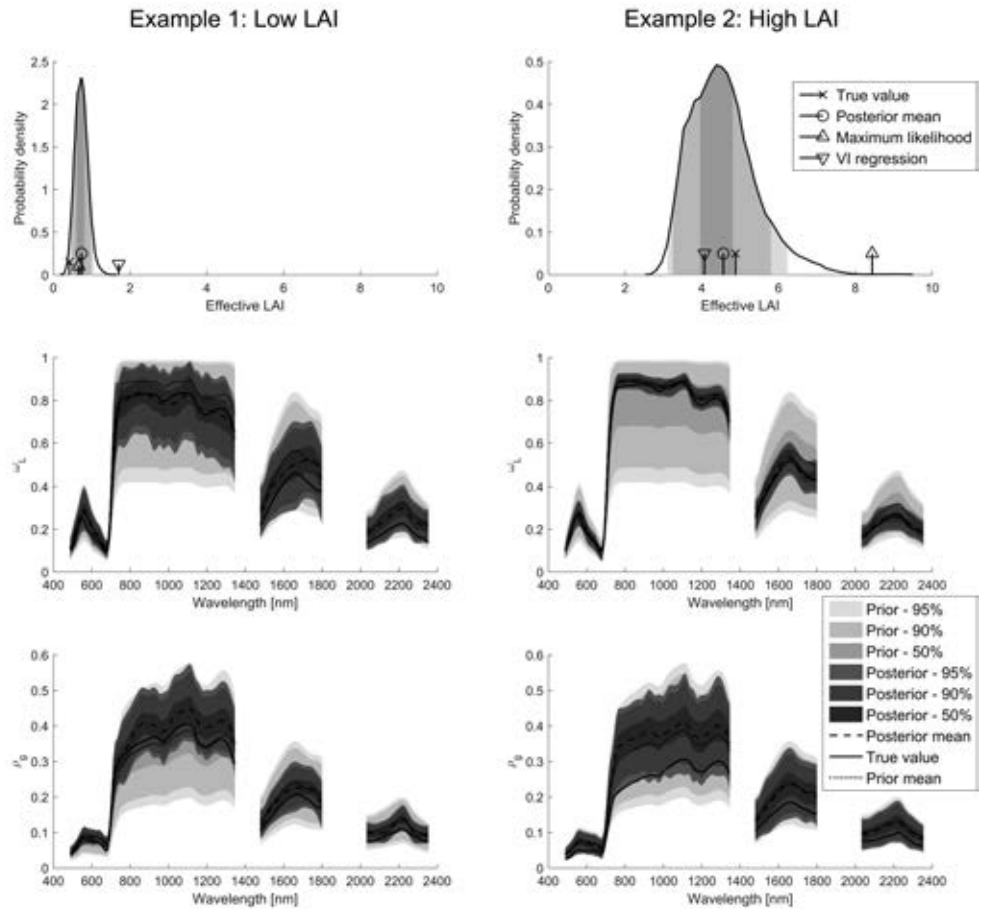

Figure 4.1: Posterior marginal densities of effective LAI (top), leaf albedo (center), and understory reflectance (bottom) for the two example stands. The shaded areas in the top figures correspond to the 50%, 90%, and 95% posterior CIs from dark to light grey, respectively.

center), the estimate values are closely aligned with the true simulation values, except in the case of few outliers. The estimation error increases somewhat with increasing LAI. However, when population mean values are used in the ML estimate (bottom right), estimation accuracy degrades tremendously. Most of the deciduous stands lie near zero and large number of spruce stands are pushed to the upper boundary at $\mathrm{LAI}_{\mathrm{eff}} = 10$. In conclusion, the ML estimate is highly sensitive to the misspecification of the secondary model parameters.

The Bayesian conditional mean estimates using the informative prior presented in Section 4.2.1 are shown in upper left in Figure 4.2. The estimates are well in line with the true simulation values. Estimation error increases moderately with increasing LAI and there is some positive bias at low LAI values (between $\mathrm{LAI}_{\mathrm{eff}} = 0$ and 1). Species-specific behavior is observed: there is a general positive bias on spruce-dominated stands, and several deciduous stands show up as clear outliers. The upper right of Figure 4.2 shows the Bayesian estimates using the uniform prior.

The estimates are highly scattered and there is a significant underestimation on many high LAI stands and a positive bias on low LAI stands that is stronger than in the informative prior results. In general, the uniform prior results are considerably worse than the informative prior results. The final reference method was the empirical VI regression (bottom left). All in all, the VI regression results are very similar to the results obtained by using the uniform prior: there is large positive bias on low LAI stands and large negative bias on high LAI stands. The high LAI behavior is, however, slightly better than in the uniform prior results. Of the four methods that do not assume the secondary model parameters as known, the Bayesian approach using informative prior formulation demonstrated the best overall performance.

In the publication **II**, the performance of the Bayesian approach using either the informative or uniform prior model in estimation of the model parameters other than LAI was also examined. In general, the informative prior produced vastly better estimates for $\mathrm{LAI}_{\mathrm{eff}}$, ω_L and ρ_g. For β, the improvement was smaller. The performance was also examined for sets classified by the dominant species of the simulated stand. The by-species RMSE difference for the informative prior Bayesian was fairly minor for $\mathrm{LAI}_{\mathrm{eff}}$ and ω_L. For ρ_g and β, more substantial by-species variation was observed. In bias, the variation between dominant species is large, depending on how well those stands are described by the informative prior model. For the RMSE, bias and CI coverage values, see Table 3 in the publication **II**.

4.5.2 Hyytiälä data set

The posterior marginal distributions for a broad-leaved example stand over the three months are shown in Figure 4.3. For details, see the publication **III**. In early May, the deciduous stand was not yet in the leaf out phase and thus LAI is close to zero (it should be noted that occlusion by branches and stems affects the optical field measurement). In June and July, leaves have grown and the field-measured $\mathrm{LAI}_{\mathrm{eff}}$ has increased to around 2.7. The same features observed in the simulated examples of the previous section (Figure 4.1) occur here: The width of the marginal density of $\mathrm{LAI}_{\mathrm{eff}}$ increases, the width of leaf albedo marginals decreases and the width of understory reflectance ρ_g marginals increases. The change in the posterior density of ρ_g is much stronger here than in the simulations. When interpreting the results, it should be kept in mind that the viewing geometry is different in the May data compared to the June and July measurements, see Table 2 in the publication **III**.

An example of the actual MCMC samples is shown in Figure 4.4. The figure contains scatter plots of every 100th MCMC sample for a single stand between each variable pair ($\mathrm{LAI}_{\mathrm{eff}}$, ω_L and ρ_g at $\lambda = 1033$ nm, and β). While the scatter plots are in a sense two-dimensional projections of the sample cloud, they illustrate the often complicated shape of the posterior density.

Scatter plots of the Hyytiälä results (publication **III**) for the Bayesian approach and the two VI regressions are presented in Figure 4.5. The Bayesian estimates (left) are more scattered than the VI regression results (center and right), but show low bias. The VI regression using the field-measured stands as a training data shows the best performance. The simulation-based VI regression has significant underestimation of $\mathrm{LAI}_{\mathrm{eff}}$.

The RMSE, relative RMSE and relative bias for the estimates are listed in Table 4.1 for all stands and grouping by month or dominant species. In the aggregated results, the VI regression using field-measured training data has the lowest RMSE and bias. The Bayesian posterior mean estimates come second, with larger RMSE and bias. The

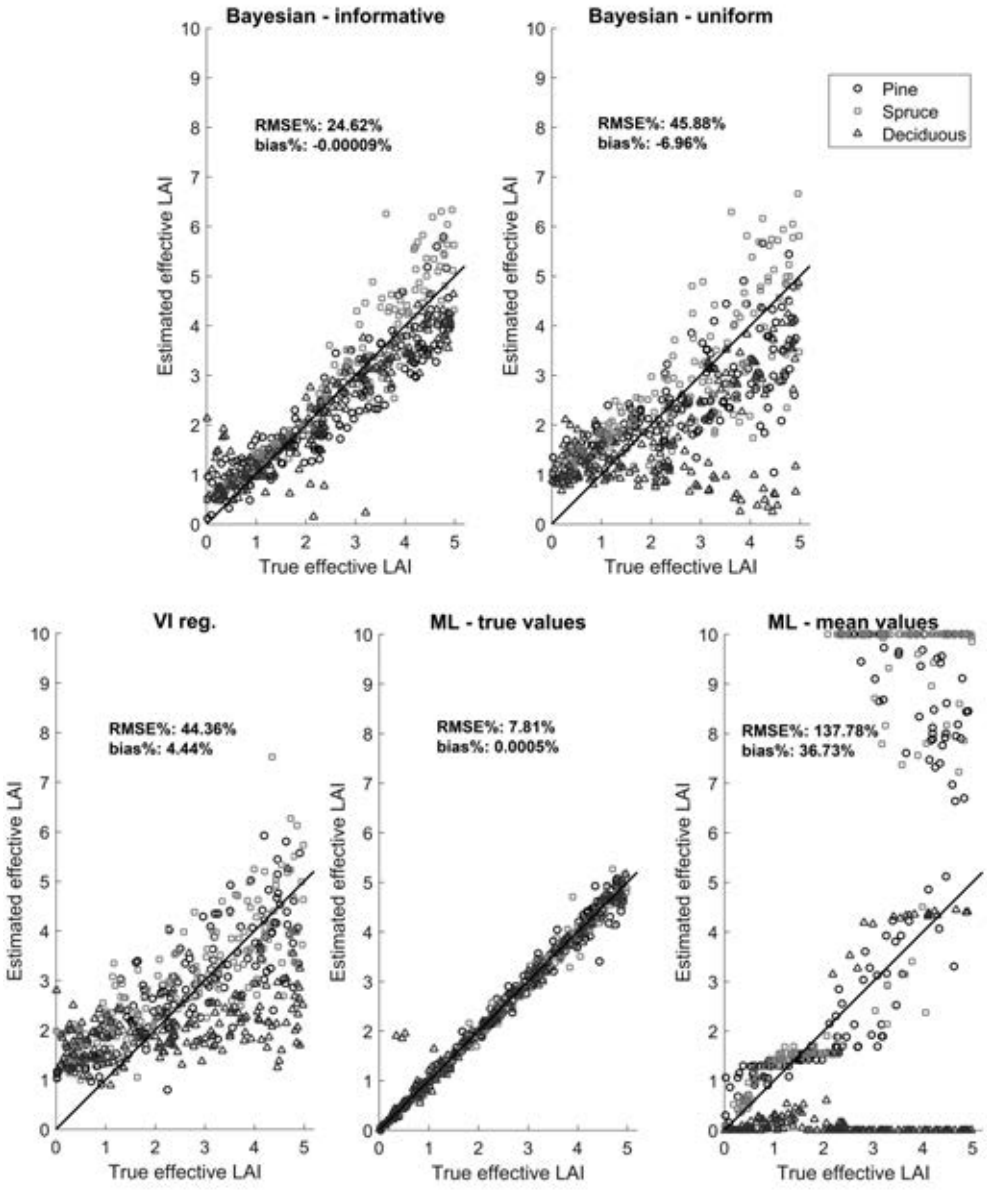

Figure 4.2: Estimated LAI_{eff} vs. true LAI_{eff} for the tested estimation methods. Pine dominated stands are marked with circles, spruce dominated with squares and deciduous with triangles. Figure also includes the relative RMSE and bias for the estimates.

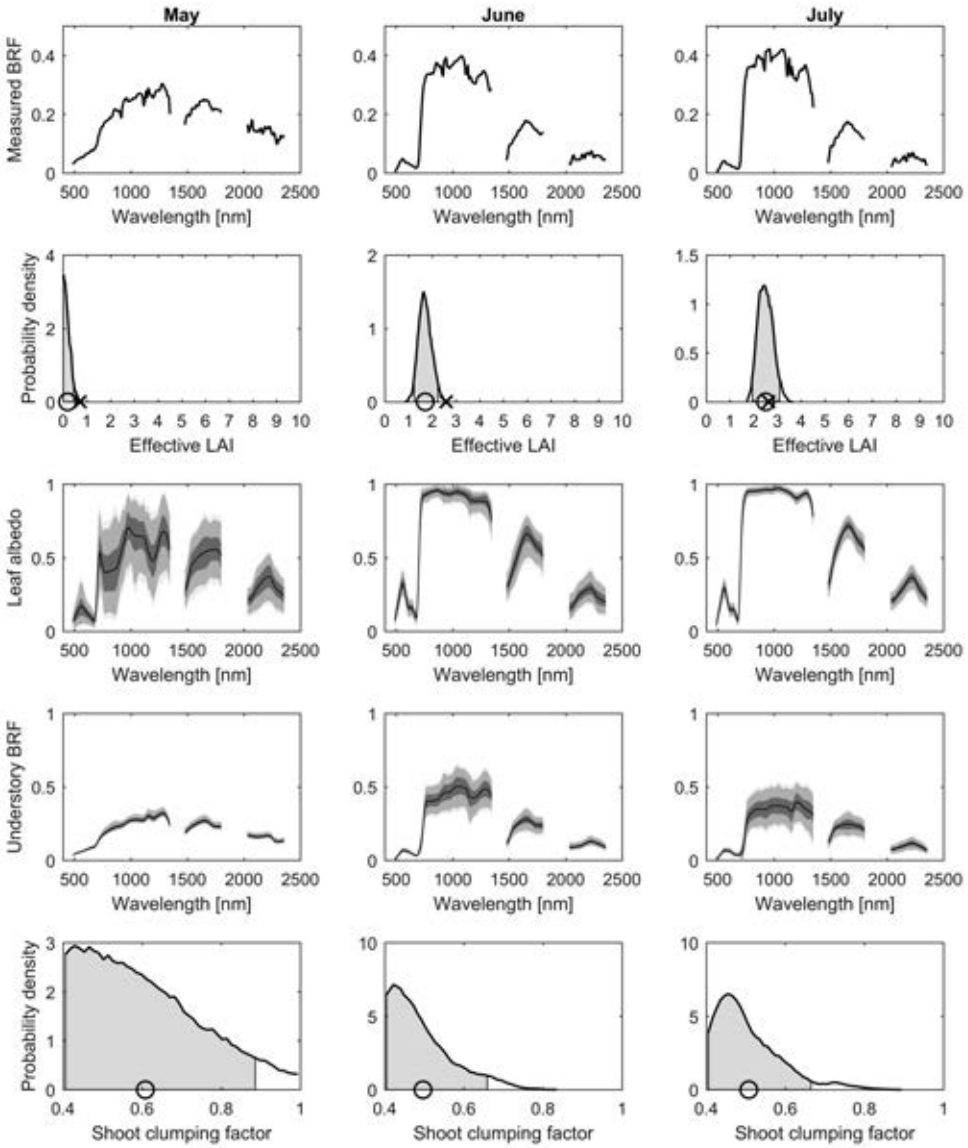

Figure 4.3: Seasonal course of a broad leaved example stand. From top to bottom: 1) The measured forest reflectance spectrum. 2) Posterior marginal density of $\mathrm{LAI}_{\mathrm{eff}}$, posterior mean is marked with a circle, the field-measured value by a cross; 95% CI is shaded. 3) Estimated leaf albedo ω_L (black line), with 50%, 90% and 95% credible envelopes (shaded area from the darkest to the lightest). 4) Estimated understory BRF ρ_g (black line), with 50%, 90% and 95% credible envelopes (shaded area from the darkest to the lightest). 5) Posterior marginal density of β, posterior mean is marked with a circle, 95% CI is shaded.

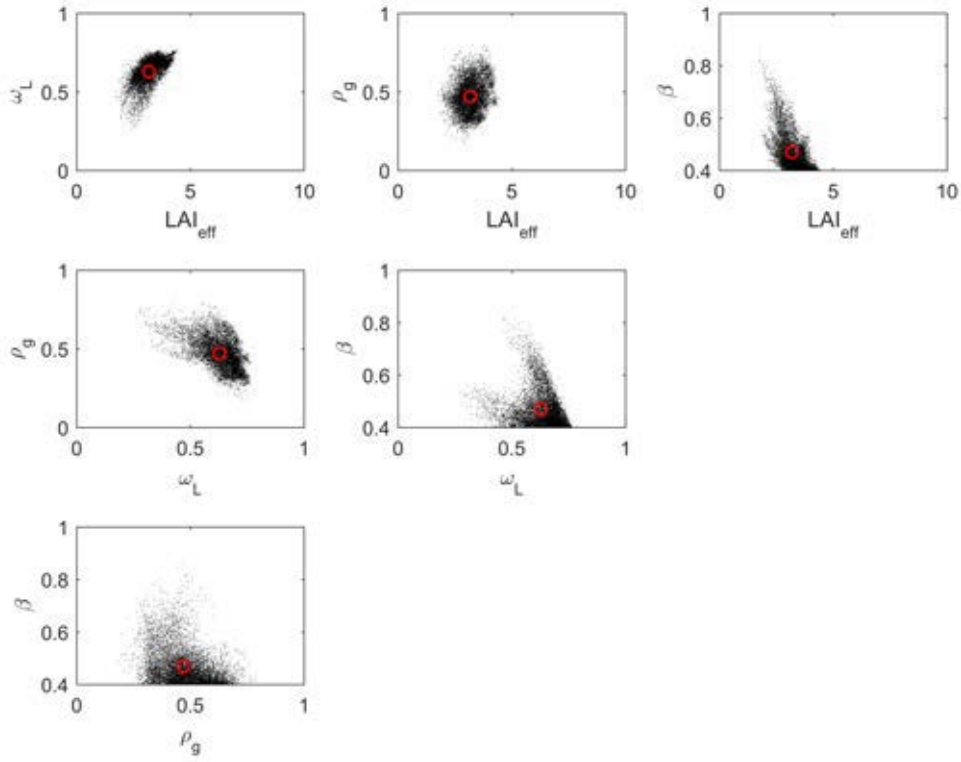

Figure 4.4: Scatter plots of the MCMC samples (every 100th sample drawn) for each variable pairs. Conditional mean estimate shown with red circle. The wavelength for ω_L and ρ_g is 1033 nm.

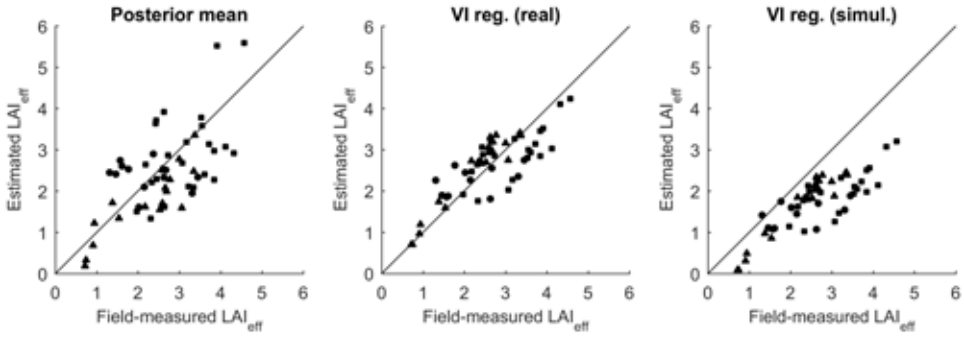

Figure 4.5: Estimated LAI_{eff} vs. true LAI_{eff} for the tested estimation methods. Pine dominated stands are marked with circles, spruce dominated with squares and deciduous with triangles.

simulation-based VI regression has a large negative bias, as was already observed in Figure 4.5. In the by-month grouping, the RMSE and bias of the VI regression estimates do not show significant monthly variation. The Bayesian estimates, on the other hand, perform well with the May measurements. In the other months, RMSE performance is only slightly better than the simulation based VI regression. In the by-species grouping, the difference between pine dominated and spruce dominated stands is insignificant for the Bayesian and the real training data based VI regression. The simulation based VI regression performs worse in spruce stands (in terms of the raw RMSE). The methods perform best in deciduous stands when measured by RMSE.

The credible interval coverage is also reported for the Bayesian estimates in Table 4.1. The CI coverage for all stands is 53.70%, which implies that the estimated intervals are significantly too narrow. The monthly and species-wise coverages range between 40% and 70%. The groupings with the best coverage values, May and Deciduous, correspond to the groups that also have the lowest RMSE values.

The estimated understory reflectance was compared with the field-measured ρ_g values. One-to-one comparison was not possible, because the field measurements of ρ_g were not done separately on every stand and have a different spectral range. However, the understory type of each stand is known and thus the ρ_g estimates corresponding to each understory grouping were compared with the field-measurements on their common shared spectral range (c. 490 – 1080 nm). The results are shown in Figure 4.6. The performance of the estimated ρ_g is fairly lackluster. The results for xeric stands in May, mesic stands in June, and herb-rich and subxeric stands in July are close to the field-measured values. Yet, for most cases there is a tendency to overestimate the ρ_g, and to produce spectra that have a stronger charasteristic structure of green vegation (e.g. prominent red edge). The analysis is somewhat confounded by the effect of canopy density (i.e. LAI) to the identifiability of ρ_g, described previously in this chapter.

4.6 DISCUSSION

Given the results, the Bayesian forest reflectance model inversion framework is suitable for estimation of canopy LAI and other forest parameters. However, there are aspects that could be feasibly improved further. Most important of these are the error model, which currently does not include any information on the model error, and the prior density.

The importance of model error (discrepancy between the model and reality) is evident when the simulation results of the publication **II** are compared with the real data results of the publication **III**. The simulated data contain no model errors and the Bayesian conditional mean estimates are practically unbiased. This unbiasedness is expected, because the conditional mean estimate is unbiased (by the law of total expectation) when the posterior density is correctly formulated. On the other hand, the real data results show significant bias. In addition to the model error, misspecification of the prior density is another possible source for the observed bias.

Ways to mitigate the model error include improving the reflectance model and statistical modeling of the model error (for example, as was done in the publication **I**). The reflectance model used in this work is fairly simple and could possibly be improved by modeling of tree bark and branches [68] and by modeling better the angular behavior of canopy scattering. However, a more complex reflectance model

Table 4.1: RMSE, relative RMSE and relative bias of effective LAI estimates for the Bayesian posterior mean and the reference VI regression estimates by month. CI coverage is also reported for the Bayesian results.

	RMSE	RMSE%	bias%	CI%
All				
Post. mean	0.83	31.77	-7.42	53.70
VI (real)	0.52	19.77	-0.57	
VI (simul.)	1.00	38.19	-32.79	
May				
Post. mean	0.59	29.88	8.66	66.67
VI (real)	0.50	24.99	-7.07	
VI (simul.)	1.03	51.70	-45.35	
June				
Post. mean	0.97	34.04	-24.58	38.89
VI (real)	0.51	17.86	0.81	
VI (simul.)	1.00	35.47	-30.56	
July				
Post. mean	0.90	29.33	-1.93	55.56
VI (real)	0.55	18.09	3.79	
VI (simul.)	0.98	31.93	-26.71	
Pine				
Post. mean	0.86	39.03	7.98	41.67
VI (real)	0.60	27.52	-4.02	
VI (simul.)	0.94	42.69	-33.34	
Spruce				
Post. mean	0.93	30.18	-2.93	50.00
VI (real)	0.58	18.31	-6.90	
VI (simul.)	1.23	38.79	-35.19	
Deciduous				
Post. mean	0.71	29.40	-21.19	66.67
VI (real)	0.35	15.67	10.47	
VI (simul.)	0.65	29.72	-27.84	

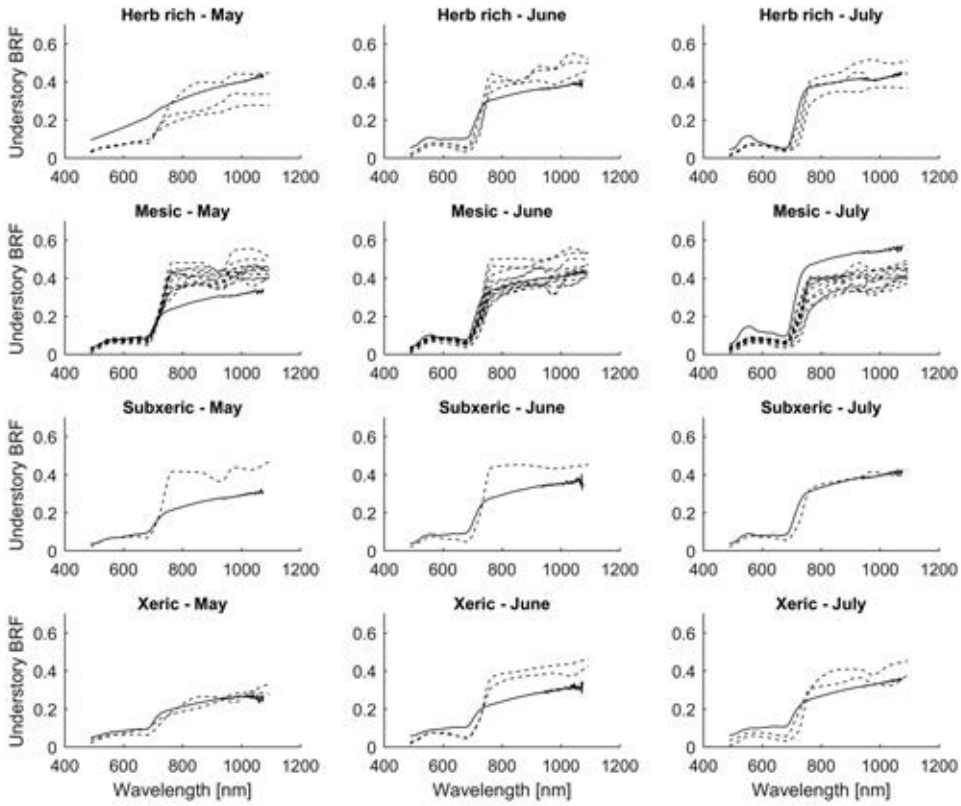

Figure 4.6: The field-measured understory reflectance (solid line) and the estimated understory reflectance (dashed line) on the common spectral interval by month and understory type. The estimated understory reflectance was computed as an average over the stand-wise estimates with the given month and understory type.

nearly always means more unknown model parameters, for example bark single scattering albedo and bark area index in the case of tree bark modeling. Statistical modeling of the model error would have negligible impact on the complexity of the inverse problem. The main problem there is the availability of suitable data for learning the error model. Probably the most feasible approach would be to use training data simulated using a complex 3D forest radiative transfer model and use the approximation error method [4, 47]. Improved error model would likely increase estimation accuracy and CI coverage.

The prior density was constructed using available data on the variables of interest. These data were measured relatively late in the growing season [65, 66], and it was not clear if the prior density was suitable for the May data. However, by the results the prior formulation works sufficiently well, yet possibly a time-dependent seasonal prior formulation could be improve the results. Another important prior improvement would be the modeling of prior correlations between the reflectance model parameters, which are currently modeled as statistically independent. However, both of these potential improvements require further empirical data on these variables

and their seasonal course.

While this research concentrated of hyperspectral data, the Bayesian framework could be just as well applied to multispectral data (e.g Landsat, Sentinel 2). Using a different instrument would require that the prior models for the spectral variables are rewritten to correspond to the spectral bands of the instrument.

5 Conclusions

In this thesis, Bayesian estimation and uncertainty quantification methods were developed for two remote sensing problems: 1) Estimation of forest inventory attributes, such as stem volume, from airborne laser scanning (ALS) data. 2) Estimation of forest canopy leaf area index (LAI) from hyperspectral satellite images. In both studied problems the developed methods were tested using real, high quality validation data.

In the publication **I**, we proposed a Bayesian inference based method for estimating species-specific stem volumes from ALS data within the so-called area-based approach. Based on a field-measured set of training plots, we constructed an approximative model connecting the stem volumes and predictors computed from the remote sensed data. This model was used for constructing a likelihood model which together with a natural prior information on the positivity of the stem volume formed the posterior distribution, which was explored with an MCMC-method to form practically useful point and interval estimates. In Chapter 3, the method was extended to estimation of multiple other species-specific stand attributes in addition to the stem volume. Furthermode, we showed how the uncertainty in the stand attribute estimates can be used to quantify uncertainty in classification done using the stand attributes.

The stem volume results of the publication **I** demonstrated that in the sense of RMSE, the point estimates produced by Bayesian approach are competitive to existing state-of-the-art methods used in area-based forest inventory. The quality of the Bayesian 95% credible intervals depended on the number of training plots used, but in the leave-one-out cross-validation the CIs were close to optimal. In the extended studies, the number of estimated stand attributes was increased, which resulted in slightly better estimation accuracy, but poorer uncertainty quantification performance. The results support the feasibility of Bayesian inference for quantifying the uncertainty of the stand attribute estimates based on ALS data.

In the publication **II**, we presented a method based on Bayesian inversion of a physically-based forest reflectance model for estimation of forest canopy LAI. Using realistic simulated data, the effect of uncertainties in the reflectance model parameters on the estimates was examined. Furthermore, we investigated if LAI estimates can recover from these uncertainties with the aid of Bayesian inference. The results of the simulation study **II** show that in the presence of unknown model parameters, the Bayesian LAI estimates which account for the model uncertainties outperform the conventional estimates that are based on biased model parameters. Moreover, the results demonstrate that the Bayesian inference can also provide feasible measures for the uncertainty of the estimated LAI. The effect of prior model formulation for the model unknowns was also tested; i.e., the informative prior formulation was compared with simple uniform prior formulation. With the informative priors the Bayesian estimates produced significantly smaller estimation errors and better estimates for the parameter uncertainty than with uniform priors.

The Bayesian inversion approach was further validated using EO-1 Hyperion satellite data in a heterogeneous Finnish boreal forest in the publication **III**. The Bayesian estimates were compared to a conventional VI regression using both field-

measured or simulated training data. Moreover, the performance of the uncertainty estimates (95% credible interval) produced by the Bayesian method was studied, and the seasonal behavior of the estimated LAI, leaf albedo and understory reflectance was evaluated. The performance of the Bayesian $\mathrm{LAI}_{\mathrm{eff}}$ estimates was superior in both RMSE and bias to the comparable simulation based VI regression. VI regression using field-measured training data was superior to both methods, but has the significant drawback of requiring field measurements. The $\mathrm{LAI}_{\mathrm{eff}}$ credible interval quality was generally fairly poor. Seasonality of the estimated leaf albedo and understory reflectance was examined. The ω_L estimates were feasible and showed a tendency to increase over the growing season. The ρ_g estimates were compared to monthly field-measurements grouped by understory type. The performance was not consistent, but in many cases promising. Several aspects of the simulation study **II**, such as variation of uncertainty in estimated ω_L and ρ_g with varying LAI, were here reproduced using real data.

The results show that Bayesian forest reflectance model inversion is feasible in estimation of forest canopy LAI, and other forest parameters, from hyperspectral satellite data. The method provides estimate uncertainty measures and the concurrent estimation of multiple forest parameters provides a potential wealth of information, not only in estimated values, but also in posterior covariances.

The uncertainty quantification methods for forest remote sensing developed in this thesis have many potential applications. Uncertainty metrics for the forest attribute estimates could be used as additional information in, e.g., forestry planning, measurement campaign design, and ecological risk assessment. In all these applications, treating estimates that have unequal uncertainty as equally trustworthy could produce suboptimal results. A simple example of utilizing the quantified estimate uncertainty was presented in this thesis when the species-specific stem volume estimates were used for probabilistic dominant species identification. Similar approach could be used for determining whether the forest needs thinning (e.g. [69]) or other treatments. The use of an informative prior density for species-specific stand attributes introduced in this thesis might provide a way to improve species-specific estimates in a temperate forest setting, where the number of distinct tree species is usually higher than three, if a good prior model for the species composition can be formulated.

Uncertainty in leaf area index is a significant problem in ecological and climate models. The methods for LAI uncertainty quantification presented here could be used to get more accurate error bounds for these models and thus reduce uncertainty in, for example, global warming predictions. The Bayesian approach may also prove to be valuable tool for small scale forest analysis. The results in this thesis support the notion that the Bayesian approach can track seasonal forest dynamics; could the approach also provide additional insight on other types of changes, such as when the forest is undergoing environmental stress, e.g. drought? In this thesis, two example applications of uncertainty quantification in remote sensing of boreal forests were evaluated. The methodology can be extended to other applications, sensors, and biomes.

BIBLIOGRAPHY

[1] G. B. Bonan, "Forests and climate change: forcings, feedbacks, and the climate benefits of forests," *Science* **320**(5882), 1444–1449 (2008).

[2] Y. Liu, J. Xiao, W. Ju, G. Zhu, X. Wu, W. Fan, D. Li, and Y. Zhou, "Satellite-derived LAI products exhibit large discrepancies and can lead to substantial uncertainty in simulated carbon and water fluxes," *Remote Sensing of Environment* **206**, 174 – 188 (2018).

[3] A. Kangas, R. Astrup, J. Breidenbach, J. Fridman, T. Gobakken, K. T. Korhonen, M. Maltamo, M. Nilsson, T. Nord-Larsen, E. Næsset, and H. Olsson, "Remote sensing and forest inventories in Nordic countries - roadmap for the future," *Scandinavian Journal of Forest Research* (2018).

[4] J. P. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems* (Springer, New York, 2005).

[5] D. Calvetti and E. Somersalo, *An Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing* (Springer, 2007).

[6] A. Tarantola, *Inverse problem theory and methods for model parameter estimation* (SIAM, Philadelphia, USA, 2005).

[7] S. E. Reutebuch, H.-E. Andersen, and R. J. McGaughey, "Light detection and ranging (LIDAR): an emerging tool for multiple resource inventory," *Journal of Forestry* **103**(6), 286–292 (2005).

[8] M. Maltamo, E. Næsset, and J. Vauhkonen, eds., *Forestry Applications of Airborne Laser Scanning* (Springer, 2014).

[9] J. Hyyppä, H. Hyyppä, D. Leckie, F. Gougeon, X. Yu, and M. Maltamo, "Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests," *International Journal of Remote Sensing* **29**(5), 1339–1366 (2008).

[10] R. E. McRoberts, W. B. Cohen, E. Næsset, S. V. Stehman, and E. O. Tomppo, "Using remotely sensed data to construct and assess forest attribute maps and related spatial products," *Scandinavian Journal of Forest Research* **25**(4), 340–367 (2010).

[11] J. E. Means, S. A. Acker, B. J. Fitt, M. Renslow, L. Emerson, C. J. Hendrix, et al., "Predicting forest stand characteristics with airborne scanning lidar," *Photogrammetric Engineering and Remote Sensing* **66**(11), 1367–1372 (2000).

[12] E. Næsset, "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data," *Remote Sensing of Environment* **80**(1), 88–99 (2002).

[13] M. Maltamo, J. Malinen, P. Packalen, A. Suvanto, and J. Kangas, "Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand-register data," *Canadian Journal of Forest Research* **36**(2), 426–436 (2006).

[14] S. Magnussen, E. Næsset, and T. Gobakken, "Prediction of tree-size distributions and inventory variables from cumulants of canopy height distributions," *Forestry* **86**(5), 583–595 (2013).

[15] P. Packalen, J. L. Strunk, J. A. Pitkänen, H. Temesgen, and M. Maltamo, "Edge-tree correction for predicting forest inventory attributes using area-based approach with airborne laser scanning," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **8**(3), 1274–1280 (2015).

[16] J. Hyyppä and M. Inkinen, "Detecting and estimating attributes for single trees using laser scanner," *The Photogrammetric Journal of Finland* **16**(2), 27–42 (1999).

[17] H.-E. Andersen, S. E. Reutebuch, and G. F. Schreuder, "Automated individual tree measurement through morphological analysis of a LIDAR-based canopy surface model," in *Proceeding of the 1st International Precision Forestry Symposium* (2001), pp. 11–22.

[18] Å. Persson, J. Holmgren, and U. Söderman, "Detecting and measuring individual trees using an airborne laser scanner," *Photogrammetric Engineering and Remote Sensing* **68**(9), 925–932 (2002).

[19] B. Koch, U. Heyder, and H. Weinacker, "Detection of individual tree crowns in airborne lidar data," *Photogrammetric Engineering & Remote Sensing* **72**(4), 357–363 (2006).

[20] S. C. Popescu, "Estimating biomass of individual pine trees using airborne lidar," *Biomass and Bioenergy* **31**(9), 646–655 (2007).

[21] T. Lähivaara, A. Seppänen, J. Kaipio, J. Vauhkonen, L. Korhonen, T. Tokola, and M. Maltamo, "Bayesian Approach to Tree Detection Based on Airborne Laser Scanning Data," *IEEE Transactions on Geoscience and Remote Sensing* **52**(5), 2690–2699 (2014).

[22] V. Junttila, M. Maltamo, and T. Kauranne, "Sparse Bayesian estimation of forest stand characteristics from airborne laser scanning," *Forest Science* **54**(5), 543–552 (2008).

[23] A. T. Hudak, N. L. Crookston, J. S. Evans, D. E. Hall, and M. J. Falkowski, "Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data," *Remote Sensing of Environment* **112**(5), 2232–2245 (2008).

[24] P. Packalen, A. Suvanto, and M. Maltamo, "A two stage method to estimate species-specific growing stock," *Photogrammetric Engineering and Remote Sensing* **75**(12), 1451–1460 (2009).

[25] H. Niska, J.-P. Skon, P. Packalen, T. Tokola, M. Maltamo, and M. Kolehmainen, "Neural networks for the prediction of species-specific plot volumes using airborne laser scanning and aerial photographs," *IEEE Transactions on Geoscience and Remote Sensing* **48**(3), 1076–1085 (2010).

[26] J. García-Gutiérrez, E. González-Ferreiro, D. Mateos-García, and J. C. Riquelme-Santos, "A Preliminary Study of the Suitability of Deep Learning to Improve LiDAR-Derived Biomass Estimation," in *Hybrid Artificial Intelligent Systems* (2016), pp. 588–596.

[27] A. O. Finley, S. Banerjee, B. D. Cook, and J. B. Bradford, "Hierarchical Bayesian spatial models for predicting multiple forest variables using waveform LiDAR, hyperspectral imagery, and large inventory datasets," *International Journal of Applied Earth Observation and Geoinformation* **22**, 147–160 (2013).

[28] S. Magnussen, G. Frazer, and M. Penner, "Alternative mean-squared error estimators for synthetic estimators of domain means," *Journal of Applied Statistics* **43**(14), 2550–2573 (2016).

[29] R. E. McRoberts, E. O. Tomppo, A. O. Finley, and J. Heikkinen, "Estimating areal means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery," *Remote Sensing of Environment* **111**(4), 466–480 (2007).

[30] J. Breidenbach, A. Nothdurft, and G. Kändler, "Comparison of nearest neighbour approaches for small area estimation of tree species-specific forest inventory attributes in central Europe using airborne laser scanner data," *European Journal of Forest Research* **129**(5), 833–846 (2010).

[31] J. M. Chen and T. A. Black, "Defining leaf-area index for non-flat leaves," *Plant, Cell and Environment* **15**, 421–429 (1992).

[32] I. Jonckheere, S. Fleck, K. Nackaerts, B. Muys, P. Coppin, M. Weiss, and F. Baret, "Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography," *Agricultural and Forest Meteorology* **121**(1–2), 19–35 (2004).

[33] G. B. Bonan, "Importance of leaf area index and forest type when estimating photosynthesis in boreal forests," *Remote Sensing of Environment* **43**(3), 303–314 (1993).

[34] T. N. Chase, R. A. Pielke, T. G. Kittel, R. Nemani, and S. W. Running, "Sensitivity of a general circulation model to global changes in leaf area index," *Journal of Geophysical Research: Atmospheres* **101**(D3), 7393–7408 (1996).

[35] F. Baret and G. Guyot, "Potentials and limits of vegetation indices for LAI and APAR assessment," *Remote sensing of environment* **35**(2-3), 161–173 (1991).

[36] G. Zheng and L. M. Moskal, "Retrieving leaf area index (LAI) using remote sensing: Theories, methods and sensors," *Sensors* **9**(4), 2719–2745 (2009).

[37] P. D'Odorico, A. Gonsamo, A. Damm, and M. E. Schaepman, "Experimental evaluation of Sentinel-2 spectral response functions for NDVI time-series continuity," *IEEE Transactions on Geoscience and Remote Sensing* **51**(3), 1336–1348 (2013).

[38] F. Baret and S. Buis, "Estimating Canopy Characteristics from Remote Sensing Observations: Review of Methods and Associated Problems," in *Advances in Land Remote Sensing*, S. Liang, ed. (Springer, 2008), pp. 173–201.

[39] K. M. Vanhatalo, M. Rautiainen, and P. Stenberg, "Monitoring the broadleaf fraction and canopy cover of boreal forests using spectral invariants," *Journal of Quantitative Spectroscopy and Radiative Transfer* **133**, 482–488 (2014).

[40] J. Heiskanen, M. Rautiainen, L. Korhonen, M. Mõttus, and P. Stenberg, "Retrieval of boreal forest LAI using a forest reflectance model and empirical regressions," *International Journal of Applied Earth Observation and Geoinformation* **13**, 595–606 (2011).

[41] J. Huang, Y. Zeng, A. Kuusk, L. Wu, B.and Dong, K. Mao, and J. Chen, "Inverting a forest canopy reflectance model to retrieve the overstorey and understorey leaf area index for forest stands," *International Journal of Remote Sensing* **32**(22), 7591–7611 (2011).

[42] Y. Knyazikhin, J. V. Martonchik, R. B. Myneni, D. J. Diner, and S. W. Running, "Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data," *Journal of Geophysical Research* **103**(D24), 32257–32276 (1998).

[43] A. Banskota, S. P. Serbin, R. H. Wynne, V. A. Thomas, M. J. Falkowski, N. Kayastha, J.-P. Gastellu-Etchegorry, and P. A. Townsend, "An LUT-based inversion of DART model to estimate forest LAI from hyperspectral data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **8**(6), 3147–3160 (2015).

[44] B. Combal, F. Baret, M. Weiss, A. Trubuil, D. Mace, A. Pragnere, R. Myneni, Y. Knyazikhin, and L. Wang, "Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem," *Remote Sensing of environment* **84**(1), 1–15 (2003).

[45] M. Rautiainen, P. Stenberg, T. Nilson, A. Kuusk, and H. Smolander, "Application of a forest reflectance model in estimating leaf area index of Scots pine stands using Landsat-7 ETM reflectance data," *Canadian Journal of remote sensing* **29**(3), 314–323 (2003).

[46] Q. Zhang, X. Xiao, B. Braswell, E. Linder, F. Baret, and B. Moore, "Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model," *Remote Sensing of Environment* **99**(3), 357–371 (2005).

[47] V. Kolehmainen, T. Tarvainen, S. R. Arridge, and J. P. Kaipio, "Marginalization of uninteresting distributed parameters in inverse problems – application to diffuse optical tomography," *International Journal for Uncertainty Quantification* **1**(1), 1–17 (2011).

[48] E. L. Lehmann and G. Casella, *Theory of point estimation* (Springer Science & Business Media, 2006).

[49] S. P. Fekete, J. S. Mitchell, and K. Beurer, "On the continuous Fermat-Weber problem," *Operations Research* **53**(1), 61–76 (2005).

[50] S. Minsker, "Geometric median and robust estimation in Banach spaces," *Bernoulli* **21**(4), 2308–2335 (2015).

[51] E. T. Jaynes, "Confidence intervals vs Bayesian intervals," in *Foundations of probability theory, statistical inference, and statistical theories of science*, W. Harper and C. Hooker, eds. (Springer, 1976), pp. 175–257.

[52] N. Turkkan and T. Pham-Gia, "Algorithm AS 308: Highest Posterior Density Credible Region and Minimum Area Confidence Region: the Bivariate Case," *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **46**(1), 131–140 (1997).

[53] W. Gilks, S. Richardson, and D. Spiegelhalter, eds., *Markov Chain Monte Carlo in Practice* (Chapman & Hall, 1996).

[54] H. Haario, M. Laine, A. Mira, and E. Saksman, "DRAM: efficient adaptive MCMC," *Statistics and Computing* **16**(4), 339–354 (2006).

[55] P. Packalen, H. Temesgen, and M. Maltamo, "Variable selection strategies for nearest neighbor imputation methods used in remote sensing based forest inventory," *Canadian Journal of Remote Sensing* **38**(5), 557–569 (2012).

[56] A. Mäkinen, A. Kangas, and L. Mehtätalo, "Correlations, distributions, and trends in forest inventory errors and their effects on forest planning," *Canadian Journal of Forest Research* **40**(7), 1386–1396 (2010).

[57] M. Nyström, N. Lindgren, J. Wallerman, A. Grafström, A. Muszta, K. Nyström, J. Bohlin, E. Willén, J. E. Fransson, S. Ehlers, et al., "Data assimilation in forest inventory: first empirical results," *Forests* **6**(12), 4540–4557 (2015).

[58] M. Rautiainen and P. Stenberg, "Application of photon recollision probability in coniferous canopy reflectance simulations," *Remote Sensing of Environment* **96**, 98–107 (2005).

[59] Y. Knyazikhin, J. Kranigk, R. B. Myneni, O. Panfyorov, and G. Gravenhorst, "Influence of small-scale structure on radiative transfer and photosynthesis in vegetation canopies," *Journal of Geophysical Research* **103**(D6), 6133–6144 (1998).

[60] P. Stenberg, "Simple analytical formula for calculating average photon recollision probability in vegetation canopies," *Remote Sensing of Environment* **109**, 221–224 (2007).

[61] T. Manninen and P. Stenberg, "Simulation of the effect of snow covered forest floor on the total forest albedo," *Agricultural and Forest Meteorology* **149**, 303–319 (2009).

[62] M. Mõttus and P. Stenberg, "A simple parameterization of canopy reflectance using photon recollision probability," *Remote Sensing of Environment* **112**, 1545–1551 (2008).

[63] J. S. Pearlman, P. S. Barry, C. C. Segal, J. Shepanski, D. Beiso, and S. L. Carman, "Hyperion, a space-based imaging spectrometer," *IEEE Transactions on Geoscience and Remote Sensing* **41**(6), 1160–1173 (2003).

[64] M. Thérézien, S. Palmroth, R. Brady, and R. Oren, "Estimation of light interception properties of conifer shoots by an improved photographic method and a 3D model of shoot structure," *Tree Physiology* **27**(10), 1375–1387 (2007).

[65] P. Lukeš, P. Stenberg, M. Rautiainen, M. Mõttus, and K. M. Vanhatalo, "Optical properties of leaves and needles for boreal tree species in Europe," *Remote Sensing Letters* **4**(7), 667–676 (2013).

[66] J. I. Peltoniemi, J. Suomalainen, E. Puttonen, J. Näränen, and M. Rautiainen, "Reflectance properties of selected arctic-boreal land cover types: field measurements and their application in remote sensing," *Biogeosciences Discussions* **5**, 1069–1095 (2008).

[67] J. Heiskanen, M. Rautiainen, P. Stenberg, M. Mõttus, and V.-H. Vesanto, "Sensitivity of narrowband vegetation indices to boreal forest LAI, reflectance seasonality and species composition," *ISPRS Journal of Photogrammetry and Remote Sensing* **78**(0), 1 – 14 (2013).

[68] P. Stenberg, P. Lukeš, M. Rautiainen, and T. Manninen, "A new approach for simulating forest albedo based on spectral invariants," *Remote Sensing of Environment* **137**, 12–16 (2013).

[69] I. Pippuri, E. Kallio, M. Maltamo, H. Peltola, and P. Packalen, "Exploring horizontal area-based metrics to discriminate the spatial pattern of trees and need for first thinning using airborne laser scanning," *Forestry* **85**(2), 305–314 (2012).

PETRI VARVIA

Due to the vast area covered by forests globally, remote sensing is needed to monitor the state of the forests. Quantifying the uncertainty in these remote sensed estimates is crucial when the estimates are further used, for example, in forestry applications or climate models. In this thesis, Bayesian uncertainty quantification of remote sensed forest attributes is presented for two example applications: one based on airborne laser scanning and the other on hyperspectral satellite imaging.

UNIVERSITY OF
EASTERN FINLAND

uef.fi

PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND
Dissertations in Forestry and Natural Sciences

ISBN 978-952-61-2866-5
ISSN 1798-5668