Determination of structural and chemical forest attributes using hyperspectral remote sensing data – Case studies in Norway spruce forests

Betreuender:
Univ.–Prof. Dr. Joachim Hill

Berichterstattende:
Univ.–Prof. Dr. Joachim Hill
Apl.–Prof. Dr. Willy Werner

Datum der wissenschaftlichen Aussprache: 14.12.2005

Trier 2006
Table of contents

ABBREVIATIONS V

LIST OF FIGURES VII

LIST OF TABLES X

CHAPTER 1 INTRODUCTION 1
1.1 Remote sensing of forest environments 1
1.2 Hyperspectral remote sensing 4
1.3 Forest characteristics for ecosystem modeling and inventory 6
1.4 Objectives and structure of thesis 8

CHAPTER 2 FUNDAMENTALS 9
2.1 Remote sensing 9
  2.1.1 Physical principles 9
  2.1.2 Atmospheric effects 12
  2.1.3 Topographic effects 15
  2.1.4 Geometric effects 16
2.2 Spectral reflectance properties of forests 16
  2.2.1 The molecular level 17
  2.2.2 The leaf level 19
  2.2.3 Soil optical properties 23
  2.2.4 The canopy level 24
  2.2.5 Spectral reflectance properties of Norway spruce 30
  2.2.6 Age dependence of forest reflectance 31
2.3 Remote sensing of forest attributes 32
  2.3.1 Classes of forest attributes 32
  2.3.2 Attributes related to forest type 33
  2.3.3 Structural attributes 34
  2.3.4 Chemical attributes 40

CHAPTER 3 MATERIAL 46
3.1 Study sites 46
  3.1.1 Idarwald 46
  3.1.2 Gerolstein 50
3.2 Field and laboratory data 50
  3.2.1 Field data 50
  3.2.2 Chemical laboratory analysis 55
  3.2.3 Spectral reflectance measurements 56
3.3 HyMap image data 57
  3.3.1 The HyMap sensor 57
  3.3.2 The HyMap data sets 59
  3.3.3 Radiometric correction 59
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.4 Geometric correction</td>
<td>64</td>
</tr>
<tr>
<td>3.3.5 Extraction of reflectance spectra from HyMap imagery</td>
<td>65</td>
</tr>
<tr>
<td>3.4 Forest Geographic Information System</td>
<td>66</td>
</tr>
<tr>
<td><strong>CHAPTER 4 STATISTICAL AND PHYSICAL METHODS</strong></td>
<td>67</td>
</tr>
<tr>
<td>4.1 Overview</td>
<td>67</td>
</tr>
<tr>
<td>4.2 Image classification</td>
<td>68</td>
</tr>
<tr>
<td>4.2.1 Introduction</td>
<td>68</td>
</tr>
<tr>
<td>4.2.2 Spectral Angle Mapper</td>
<td>69</td>
</tr>
<tr>
<td>4.3 Statistical methods</td>
<td>70</td>
</tr>
<tr>
<td>4.3.1 Correlation analysis</td>
<td>70</td>
</tr>
<tr>
<td>4.3.2 Regression analysis</td>
<td>71</td>
</tr>
<tr>
<td>4.3.3 Artificial neural networks</td>
<td>73</td>
</tr>
<tr>
<td>4.3.4 Ratio-based and orthogonal vegetation indices</td>
<td>74</td>
</tr>
<tr>
<td>4.3.5 The red edge inflection point</td>
<td>75</td>
</tr>
<tr>
<td>4.3.6 Derivative spectra</td>
<td>77</td>
</tr>
<tr>
<td>4.3.7 Continuum-removal and band-depth normalisation</td>
<td>78</td>
</tr>
<tr>
<td>4.3.8 Non-linear least-squares spectral matching</td>
<td>79</td>
</tr>
<tr>
<td>4.4 Physically-based reflectance modeling</td>
<td>80</td>
</tr>
<tr>
<td>4.4.1 Canopy reflectance models</td>
<td>80</td>
</tr>
<tr>
<td>4.4.2 The Invertible Forest Reflectance Model</td>
<td>81</td>
</tr>
<tr>
<td>4.4.3 Inversion of reflectance models</td>
<td>83</td>
</tr>
<tr>
<td>4.5 Validation</td>
<td>84</td>
</tr>
<tr>
<td><strong>CHAPTER 5 MAPPING FOREST COVER TYPE THROUGH IMAGE CLASSIFICATION</strong></td>
<td>86</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>86</td>
</tr>
<tr>
<td>5.2 Implementation of additional information</td>
<td>87</td>
</tr>
<tr>
<td>5.3 Image classification</td>
<td>89</td>
</tr>
<tr>
<td>5.3.1 Data preprocessing</td>
<td>89</td>
</tr>
<tr>
<td>5.3.2 Training</td>
<td>90</td>
</tr>
<tr>
<td>5.3.3 Classification</td>
<td>91</td>
</tr>
<tr>
<td>5.3.4 Validation</td>
<td>92</td>
</tr>
<tr>
<td>5.4 Results and discussion</td>
<td>92</td>
</tr>
<tr>
<td><strong>CHAPTER 6 DETERMINATION OF FOREST STRUCTURAL ATTRIBUTES</strong></td>
<td>97</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>97</td>
</tr>
<tr>
<td>6.2 Developing predictive models</td>
<td>98</td>
</tr>
<tr>
<td>6.2.1 Introduction</td>
<td>98</td>
</tr>
<tr>
<td>6.2.2 Computation of vegetation indices</td>
<td>99</td>
</tr>
<tr>
<td>6.2.3 Analysis of forest stand variables</td>
<td>99</td>
</tr>
<tr>
<td>6.2.4 Relationships between vegetation indices and forest variables</td>
<td>106</td>
</tr>
<tr>
<td>6.2.5 Absorption features related to leaf area index and crown volume</td>
<td>112</td>
</tr>
<tr>
<td>6.2.6 Maps of forest attributes</td>
<td>112</td>
</tr>
<tr>
<td>6.2.7 Conclusions</td>
<td>113</td>
</tr>
<tr>
<td>6.3 Inversion of a forest reflectance model</td>
<td>118</td>
</tr>
<tr>
<td>6.3.1 Introduction</td>
<td>118</td>
</tr>
<tr>
<td>6.3.2 Parameterisation of the forest reflectance model</td>
<td>118</td>
</tr>
<tr>
<td>6.3.3 Overview of the inversion procedure</td>
<td>123</td>
</tr>
<tr>
<td>6.3.4 Forward reflectance modeling</td>
<td>123</td>
</tr>
<tr>
<td>6.3.5 Spectral sampling</td>
<td>124</td>
</tr>
</tbody>
</table>
Abbreviations

5S 5S Simulation of the Satellite Signal in the Solar Spectrum
ANN Artificial neural network
AVIRIS Airborne Visible/Infrared Imaging Spectrometer
BDN Band-depth normalisation, Band depth-normalised spectra
BLR Bivariate linear regression
C_AB Chlorophyll a+b concentration
CASI Compact Airborne Spectrographic Imager
CD Crown diameter
CET Central European Time
CF Common factor
C_N Nitrogen concentration
CO Canopy cover, fraction of canopy closure
CR Continuum-removal, Continuum-removed spectra
CR model Canopy reflectance model
C_W Water concentration
DAIS Digital Airborne Imaging Spectrometer
DEM Digital elevation model
dm dry matter
DN Digital number
DRFL First difference of reflectance spectra
DWRs First difference of water removed spectra
ENVI Environment for Visualisation of Images
F-data Field-Spec-II data
FH-data Field-Spec-II data resampled to HyMap wavebands
FHN-data Field-Spec-II data resampled to HyMap wavebands plus noise
FoGIS Forest Geographic Information System
GCP Ground Control Points
GIS Geographic Information System
GPS Global Positioning System
GVI Greeness vegetation index
H Canopy height
HyMap Hyperspectral Mapper
IGM Inverted Gauss model
INFORM Invertible forest reflectance model
LAI Leaf area index, effective leaf area index
LAI_c Canopy leaf area index
LAI_s Single tree leaf area index
LGI Lagrange interpolation
LNC Leaf nitrogen content or concentration
LSM Least squares spectral matching
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LWC</td>
<td>Leaf water content or concentration</td>
</tr>
<tr>
<td>mid-IR</td>
<td>Middle infrared</td>
</tr>
<tr>
<td>MLR</td>
<td>Multivariate linear regression</td>
</tr>
<tr>
<td>MNF</td>
<td>Minimum Noise Fraction</td>
</tr>
<tr>
<td>MVI</td>
<td>Mid-infrared vegetation index</td>
</tr>
<tr>
<td>NDVI</td>
<td>Normalised difference vegetation index</td>
</tr>
<tr>
<td>near-IR</td>
<td>Near infrared</td>
</tr>
<tr>
<td>OAA</td>
<td>Overall Accuracy</td>
</tr>
<tr>
<td>PBH</td>
<td>Perimeter-at-breast-height</td>
</tr>
<tr>
<td>PC</td>
<td>Principal component</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PLSR</td>
<td>Partial least squares regression</td>
</tr>
<tr>
<td>PVI</td>
<td>Perpendicular vegetation index</td>
</tr>
<tr>
<td>R</td>
<td>Reflectance</td>
</tr>
<tr>
<td>r</td>
<td>Correlation coefficient</td>
</tr>
<tr>
<td>R²</td>
<td>Coefficient of determination</td>
</tr>
<tr>
<td>REIP</td>
<td>Red edge inflection point</td>
</tr>
<tr>
<td>RFL</td>
<td>Reflectance spectra</td>
</tr>
<tr>
<td>RGB</td>
<td>Red-Green-Blue</td>
</tr>
<tr>
<td>rmse</td>
<td>Root mean squared error</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of interest</td>
</tr>
<tr>
<td>RVI</td>
<td>Ratio vegetation index</td>
</tr>
<tr>
<td>SAM</td>
<td>Spectral Angle Mapper</td>
</tr>
<tr>
<td>SBM</td>
<td>Stem biomass</td>
</tr>
<tr>
<td>SD</td>
<td>Stem density</td>
</tr>
<tr>
<td>SMR</td>
<td>Stepwise multiple linear regression</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td>SWIR</td>
<td>Short wave infrared</td>
</tr>
<tr>
<td>thermal-IR</td>
<td>Thermal infrared</td>
</tr>
<tr>
<td>TM</td>
<td>Thematic Mapper</td>
</tr>
<tr>
<td>TSAVI</td>
<td>Transformed soil-adjusted vegetation index</td>
</tr>
<tr>
<td>VI, VIs</td>
<td>Vegetation index, Vegetation indices</td>
</tr>
<tr>
<td>VOL</td>
<td>Crown volume</td>
</tr>
<tr>
<td>WRS</td>
<td>Water removed spectra</td>
</tr>
</tbody>
</table>
List of figures

Figure 1.1: Vegetation parameters in forests with relevance to ecosystem simulation models .............................................................................................................. 2
Figure 1.2: Strategies to extract forest attributes from remote sensing data ...................... 3
Figure 1.3: The concept of imaging spectrometry .............................................................................................................................. 5
Figure 2.1: Interactions of incoming solar radiation with the atmosphere and the contamination of the satellite-measured radiance ................................................................................................. 13
Figure 2.2: Absorption features of organic materials and water ........................................ 17
Figure 2.3: Typical spectra of leaf reflectance, transmittance and absorptance ................. 20
Figure 2.4: Internal structure of a dicotyledonous leaf (from Govaerts et al., 1996) and interaction of leaf tissues with incoming radiation ................................................................................................. 21
Figure 2.5: Spectrum of specific absorption coefficients of \textit{in vivo} chlorophyll a+b and \textit{in vivo} carotenoids ................................................................................................................................. 22
Figure 2.6: Classification of soil spectral curves into five general types ............................... 23
Figure 2.7: Schematic representation of the changes in canopy reflectance during growth and senescence ........................................................................................................................................... 24
Figure 2.8: Schematic variation of canopy reflectance as a function of leaf area index for the red and near-IR domain ........................................................................................................................................... 25
Figure 2.9: Oblique (left) and nadir (right) view of black spruce (\textit{Picea mariana}) demonstrating the gap effect ........................................................................................................................................... 26
Figure 2.10: Black spruce (\textit{Picea mariana}) observed in backward scattering (left) and forward scattering (right) direction ........................................................................................................................................... 26
Figure 2.11: Causes for directional reflectance anisotropy of vegetation canopies and the associated typical shapes of the BRDF on the principal solar plane ........................................................................................................................................... 27
Figure 2.12: Variation of simulated forest canopy reflectance .................................................. 29
Figure 2.13: Measured reflectance spectra of \textit{Norway spruce} at the needle, branch, and canopy level ......................................................................................................................................................... 30
Figure 2.14: Measured reflectance spectra of \textit{Norway spruce} for changing numbers of branches and corresponding values of cumulative LAI ........................................................................................................................................... 31
Figure 2.15: Canopy reflectance of \textit{Norway spruce} as a function of stand age ................. 32
Figure 3.1: Location of study sites .......................................................................................... 47
Figure 3.2: General map of Idarwald test site ...................................................................... 51
Figure 3.3: General map of Gerolstein test site with stand boundaries, locations of field plots and locations of targets used for radiometric calibration ................................................................. 53
Figure 3.4: Woodman climbing a tree at Idarwald test site .................................................... 54
Figure 3.5: Signal-to-noise ratio measured while viewing a 50% reflectance target ............ 56
Figure 3.6: Raw (left) and cross-track illumination corrected (right) HyMap image profiles (upper part) and corresponding portions of the HyMap scene (lower part) at Idarwald ........................................................................................................................................... 58
Figure 3.7: Schematic representation of the sensor recalibration ........................................... 59
Figure 3.8: Atmospherically corrected HyMap reflectance spectra from various forested and non-forested areas at Idarwald study site ........................................................................................................... 60
Figure 3.9: Modelled against measured radiance in six wavebands based on ten reference targets at Gerolstein test site .......................................................... 63
Figure 3.10: Extraction of spectra for various radii .......................................................................................................................... 64
Figure 3.11: (A) Change of reflectance with increasing circle radius across the visible region (HyMap-bands 4-12) and (B) the near-infrared region (HyMap-bands 20-32). (C) Change of the coefficient of variation with increasing circle radius across the visible region (HyMap-bands 4-12) and (D) the near-infrared region (HyMap-bands 20-32) ..................................................................................... 65
Figure 4.1: Continuum-removal for chlorophyll absorption in vegetation .......................................................... 78
Figure 4.2: Schematic view of the link between LIBERTY, SAIL, and FLIM models showing the input variables used for the simulation .......................................................................................................................... 80
Figure 5.1: Four datasets of varying spectral and spatial resolution .......................................................................................................................... 87
Figure 5.2: Estimated against field-measured stem density (cross-validated) at Idarwald test site ........................................................................................................... 88
Figure 5.3: Stem density (unit: ha⁻¹) at the Idarwald obtained from black/white orthophotos .......................................................................................................................... 88
Figure 5.4: Reflectance spectra of forest types considered in the classification process .......................................................................................................................... 90
Figure 5.5: Map of forest type at Idarwald .......................................................................................................................... 96
Figure 5.6: LAI as a function of stand age for Norway spruce stands at Idarwald test site .......................................................................................................................... 100
Figure 6.2: Correlograms for six stand attributes and both reflectances (left) and first derivative reflectance spectra (right) .......................................................................................... 102
Figure 6.3: Crown volume (VOL) as a function of stand age for Norway spruce stands at Idarwald test site .......................................................................................................................... 104
Figure 6.4: Stem biomass (SBM) as a function of stand age for Norway spruce stands at Idarwald test site .......................................................................................................................... 104
Figure 6.5: 2D-correlation plot that shows the correlation (R²) between LAI and narrow band RVI values .......................................................................................................................... 106
Figure 6.6: Linear regression between best narrow band PVI and LAI .......................................................................................................................... 110
Figure 6.7: Linear regression between best narrow band PVI and crown volume .......................................................................................................................... 111
Figure 6.8: Wavelength position of selected optimal narrowband VIs for LAI (left) and VOL (right) .......................................................................................................................... 112
Figure 6.9: Map of effective leaf area index at Idarwald obtained from an empirically derived predictive equation .......................................................................................................................... 115
Figure 6.10: Map of crown volume at Idarwald obtained from an empirically derived predictive equation .......................................................................................................................... 116
Figure 6.11: Map of stand mean leaf area index at Idarwald .......................................................................................................................... 117
Figure 6.12: Measured (red) and modelled (blue) forest canopy reflectance for certain values of canopy LAI .......................................................................................................................... 121
Figure 6.13: Measured (red) and modelled (blue) forest canopy reflectance against canopy LAI .......................................................................................................................... 121
Figure 6.14: Modelled versus measured forest canopy reflectance for certain values of canopy LAI in different wavebands .......................................................................................................................... 122
Figure 6.15: General procedure for estimating canopy attributes from remote sensing data through inversion of a forest reflectance model .......................................................................................................................... 123
Figure 6.16: Structure of the Artificial Neural Network used to invert the INFORM model .......................................................................................................................... 125
Figure 6.17: Modelled against measured forest canopy variables .......................................................................................................................... 126
Figure 6.18: Map of effective leaf area index at Idarwald obtained through inversion of a forest reflectance model ................................................................. 128
Figure 6.19: Map of stem density at Idarwald obtained through inversion of a forest reflectance model ................................................................. 129
Figure 7.1: Location of absorption features related to total chlorophyll (a), water (b), and protein (c, d, e) used in the regression analysis .................................................. 133
Figure 7.2: Cross-validated rmse against number of wavebands selected by stepwise multiple regression for the Field-Spec-II data resampled to HyMap spectral bands (DWRS data set) .......................................................... 134
Figure 7.3: A) Matching of fresh needle (stack layer) spectrum in the 1500-1780 nm region with the water spectrum and with a combination of the water and the protein spectrum, B) residual spectra between the measured and the modeled spectra ........................................................................................................ 136
Figure 7.4: Positions of wavebands that have been selected by SMR between DWRS spectral data (resampled to HyMap wavebands) and three leaf chemicals. 139
Figure 7.5: Estimated against measured leaf total chlorophyll concentration ($C_{AB}$) (left) and estimated against measured leaf nitrogen concentration ($C_N$) (right) .... 150
Figure 7.6: Map of leaf chlorophyll concentration at Gerolstein test site ...................... 151
Figure 7.7: Map of leaf nitrogen concentration at Gerolstein test site ......................... 152
Figure 7.8: Statistical measures for three main phases and five main geological epochs that occur in the study area ................................................................. 153
List of tables

Table 2.1: Absorption features in vegetation reflectance spectra related to leaf nitrogen concentration.................................................................................................. 19
Table 2.2: Selected literature results on the estimation of forest structural attributes. .... 37
Table 2.3: Selected literature results on the estimation of foliar chemicals from spectral variables using different methodologies ......................................................... 41
Table 3.1: Soil substrates occurring at the Gerolstein test site ....................................... 48
Table 3.2: Soil chemical indicators for three typical soil profiles occurring at the Gerolstein test site.......................................................................................... 49
Table 3.3: Spectral configuration of the HyMap sensor................................................... 57
Table 4.1: Broadband vegetation indices investigated in this study ................................ 75
Table 5.1: Training data used in the forest classification................................................. 89
Table 5.2: Summary of accuracies.................................................................................. 91
Table 5.3: Confusion matrices for 5 different datasets: A: hyperspectral, 5 m; B: multispectral, 5 m; C: hyperspectral, 30 m; D: multispectral 30 m, E: hyperspectral 5 m and orthophoto derived stem density................................ 93
Table 6.1: Resulting subsets after stratification according to forest stand age ............... 97
Table 6.2: Summary statistics for forest stand attributes................................................. 98
Table 6.3: Linear correlation between forest stand variables ........................................... 98
Table 6.4: Linear correlation between forest stand attributes and HyMap reflectance resampled to TM spectral bands ................................................................. 100
Table 6.5: Statistics of the first component factor of the varimax rotation (CF1) compared to the Eigenvector of the first principal component PC1........................................ 103
Table 6.6: Linear correlation between principal components and HyMap reflectance resampled to TM spectral bands ................................................................. 103
Table 6.7: Best narrow band RVI and PVI derived from 2D-correlation plots for different subsets ......................................................................................................... 107
Table 6.8: Cross-validated R² (first line) and cross-validated rmse (second line) for linear regression between broadband and hyperspectral (narrowband VIs and REIP) indices and forest stand variables LAI and VOL........................................ 109
Table 6.9: Summary statistics for leaf area index in five age classes ......................... 113
Table 6.10: Relationships between canopy leaf area index and the attributes canopy cover, stem density and canopy height ............................................................... 119
Table 6.11: Inputs to the INFORM model used to compare modelled with measured reflectance spectra derived from field data and modelled values of CO and LAIc ................................................................. 119
Table 6.12: Important parameters that were kept constant during all model simulations......................................................................................................................... 120
Table 6.13: Statistics of canopy variables used to simulate forest reflectance .......... 124
Table 7.1: Summary of data for three chemicals at two sites. $C_{AB}$ = Chlorophyll a+b concentration, $C_N$ = Nitrogen concentration, $C_W$ = Water concentration ...... 132
Table 7.2: Five absorption features used in the regression analysis.......................... 133
Table 7.3: The cross-validated coefficient of determination ($R^2$) and the cross-validated rmse between two sets of spectral data (features at 1800 nm, 2100 nm, and 2300 nm) and nitrogen and water for 1st year needle reflectance spectra resampled to HyMap spectral bands ............................................................ 137

Table 7.4: The cross-validated coefficient of determination ($R^2$) and the cross-validated rmse between spectral data (DWRS) and concentrations of chlorophyll, nitrogen and water........................................................................................ 138

Table 7.5: Wavebands that have been selected by SMR between DWRS spectral data (resampled to HyMap wavebands) and three leaf chemicals ................. 140

Table 7.6: Wavelengths correlated with nitrogen (Nit), protein (Pro), total chlorophyll (Chl) and water (Wat) by multivariate statistical measures ................................. 142

Table 7.7: The cross-validated coefficient of determination ($R^2$) and the cross-validated rmse between three types of spectral data (resampled to HyMap wavebands) and three leaf chemicals.................................................... 143

Table 7.8: Wavebands selected by SMR using three sets of spectral data (resampled to HyMap wavebands) and three leaf chemicals........................................... 143

Table 7.9: The performance of regression equations developed from the calibration dataset (Gerolstein) that were afterwards applied to the validation data set (Idarwald) for estimation of chemicals .......................................................... 145

Table 7.10: The cross-validated coefficient of determination ($R^2$) and the cross-validated rmse between five types of spectral data (HyMap imagery) and three leaf chemicals .............................................................................. 147

Table 7.11: Wavebands selected by SMR using five sets of spectral data (HyMap imagery) and three leaf chemicals ................................................................. 148

Table 7.12: Wavebands that have been selected by SMR between CR spectral data (HyMap imagery) and three leaf chemicals ...................................................... 149

Table 8.1: Determination of leaf area index ($n = 39$) using different methods and data .................................................................................................................. 159

Table 8.2: Determination of chemical forest attributes using laboratory and HyMap data and three types of spectral derivatives ............................................. 160
Chapter 1 Introduction

1.1 Remote sensing of forest environments

A significant progress has been realised in the remote sensing of forests in recent years linked to technological advances in sensor design, growth in information extraction techniques and increasing requirement to quantitatively describe and understand our environment (Wulder & Franklin, 2003). Remote sensing instruments measure the radiation that is emitted or reflected from the earth’s surface. The diversity of instruments available at present and in the future provide data with broad to fine spectral resolution, with large to small spatial resolution and other characteristics (e.g., multi-directionality) suitable for the quantitative and qualitative analysis of forests (Peterson & Running, 1989). A relatively new optical sensor type, the imaging spectrometer, represents a technological advance in remote sensing, as it measures the reflected signal with a very fine spectral resolution (Curran, 1994). These hyperspectral measurements need to be transformed into vegetation attributes in order to use such data for certain applications.

As a result of the technological and methodological advances the need of information derived from forest remote sensing has risen considerably in applications such as forest inventory and management, monitoring of forest fire and disease and environmental change assessment. While traditionally forest information for inventory purposes was gathered in situ or through air photography, remote sensing is becoming central, as an increasing spatial and temporal resolution of observations is desired (Cihlar et al., 2003). The increasing demand for information on the productivity of the world’s forests has led to new approaches beyond traditional ground-based forest inventory techniques. The development of physiologically-based process models together with digital remote sensing methodologies allow for a better prediction of forest growth patterns across landscapes (Coops & White, 2003). Remotely sensed forest attributes may be used to initialise, update, and validate ecosystem simulation models. As a key variable, leaf area index is often used to couple remote sensing with process models. The determination of forest attributes by means of remote sensing requires the usage of appropriate sensors and methodologies. Current concepts to remotely sense ecosystem-relevant forest attributes are schematically illustrated in Figure 1.1.
While previously pre-processing of remote sensing images, such as geometric and radiometric adjustments, occupied a large portion of the information extraction process, image analysis has become increasingly able to process the available data into the actual information desired (Wulder & Franklin, 2003). Information extraction from remote sensing data can be described in terms of models of the sensor, atmosphere, and the scene (Strahler et al., 1986; Franklin et al., 2003): In remote sensing of forests, the scene comprises a forested portion of the earth surface viewed at a specific scale. An appropriate image processing strategy is applied depending on whether the attribute to be mapped occurs over spatial extents that are larger or smaller than the spatial resolution of the sensor (Figure 1.2). Mapping of the attribute ‘forest type’ (e.g. tree species, age class) assumes that the image pixel is completely covered by one forest type and thus, many pixels form a forest stand. In this case, image classification would be a suitable processing algorithm. Multispectral classification techniques are used to map

Figure 1.1: Vegetation parameters in forests with relevance to ecosystem simulation models, i.e. biogeochemical (BGC) models, that are open to remote sensing include cover type, leaf area index (LAI), and leaf nitrogen content (LNC). Cover type is usually mapped at the landscape or canopy level using image classification techniques of multispectral satellite imagery. LAI and other structural attributes can be estimated through statistical or physical modeling approaches where both multi- and hyperspectral data may be implemented. Efforts to determine leaf chemical attributes, such as LNC or chlorophyll content commonly start at the leaf level and try to progress towards the canopy and landscape levels.
qualitative attributes into distinct classes. When the focus of the analysis is on attributes that vary within a stand on a sub-pixel level (e.g. leaf area index, canopy cover, biomass or chlorophyll concentration), quantitative processing algorithms have to be chosen. Quantitative attributes can generally be derived through statistical (empirical) or physical approaches. The empirical approach relates measurements of reflectance or vegetation indices to the desired forest variable through a statistically-derived equation. The physical approach calculates reflectance from leaf, canopy and stand characteristics using mathematical equations based on physical principles. These reflectance models can be either used in the “forward mode” to simulate reflectance characteristics from given object properties or in the “inverse mode” to estimate the object property from given reflectance data (Goel, 1989; Asner et al., 2003). Opposed to empirical regression, physically-based approaches have the advantage that they can take into account effects of varying canopy structure, canopy shadow, understorey and soil that all influence the radiation measurements. Spectral mixture analysis uses simple reflectance models to estimate the fractional cover of major types of materials within image pixels.

Forests are generally challenging targets as a consequence of tree dimensions, architectural heterogeneity and spatial-temporal foliage dynamics. Remote sensing of forest structural and

Figure 1.2: Strategies to extract forest attributes from remote sensing data as implemented within the framework of this research. See text and section 1.4 for further details.
chemical attributes is further complicated by the contribution of understorey vegetation, litter, soil, bark as well as plant and relief shadow, all of which influence the whole radiometric signal (Spanner et al., 1990; Chen et al., 1999; Soudani et al., 2002). Mapping of forest cover becomes a particularly difficult task when trees of different species and age form mixed stands. Managed forests, in particular coniferous stands exhibit a number of advantages to remote sensing. The forest stands have a relatively homogenous structure as trees within a single stand are of the same species and age. Coniferous trees, such as Norway spruce (Picea abies L. Karst.), with its relatively large number of foliage age classes, do not show very pronounced seasonal changes in stand structure (Vose et al., 1994) and are well suited to combine remotely sensed imagery with field-measured data.

There is a substantial interest in forest monitoring at continental, regional, and local scales. However, the present availability of hyperspectral remote sensing imagery of highest data quality (e.g., the HyMap sensor) is limited to airborne systems which collect data of limited spatial extent. Therefore, the focus of this research is on areas that can be covered by single HyMap scenes and thus, on local studies.

1.2 Hyperspectral remote sensing

Hyperspectral remote sensing or imaging spectrometry refers to the recording of remote sensing data using imaging spectrometers. These instruments collect radiance values in many narrow bands forming continuous spectra. As imaging spectrometer data is characterised by a high spectral resolution compared to multispectral data it is also called hyperspectral. The concept of imaging spectrometry is shown in Figure 1.3. Reflected solar radiation from the earth’s surface is dispersed in a spectrometer and recorded in 100-200 spectral images of the scene. For each pixel of the scene a continuous reflectance spectrum is obtained (Vane & Goetz, 1988).

Imaging spectrometry originated from the merging of spectroscopy and imaging. Reflectance spectroscopy, the study of reflected light as a function of wavelength, has been used as a quantitative tool in the laboratory for many years. Optical imaging sensors have been developed since the 1970s (e.g. Landsat MSS, Landsat TM). The first images from an airborne imaging spectrometer (AIS) were acquired in the early 1980s and up to now a wide range of systems have been developed. Several airborne hyperspectral sensors, such as AVIRIS or HyMap, have operated successfully. Spaceborne hyperspectral sensor face certain limitations compared to airborne systems: the lower signal-to-noise ratio (SNR), significantly more atmosphere effects, and detector calibration issues. Nonetheless, first space-borne sensors of more or less hyperspectral type designed for land imaging were launched during the last years. On board of the TERRA platform (launch: 1999) are the Advanced Spaceborne Thermal Emission and Reflectance Radiometer (ASTER) and the Moderate resolution
imaging spectroradiometer (MODIS). EO-1 (launch 2000) hosts the Hyperion hyperspectral imager and the Advanced Land Imager (ALI). In 2002, Envisat which carries the MERIS (Medium Resolution Imaging Spectrometer) instrument was launched by the European Space Agency. Multi-angular systems, that collect spectral data observed under different directions, help to determine structural vegetation characteristics. Examples of such systems are the hyperspectral sensor CHRIS on PROBA (launched in October 2001) and POLDER on ADEOS-2 (launch scheduled for late 2002). Despite an emerging fleet of satellites providing frequent, high spectral resolution data to users, airborne systems will continue to play an important role in forest remote sensing due to their high data quality and their flexible use.

Imaging spectrometry was used initially in geological, then in aquatic and later in ecological and atmospheric research (Curran, 1994). The motivation for the development of imaging spectrometry was mineral identification through diagnostic absorption features (Van der Meer, 1994). These narrow spectral features can not be resolved by multispectral sensors such as Landsat TM because its spectral bandwidth is too wide and the bands are not contiguous (Vane & Goetz, 1988). Plants, in contrast, are composed of the same few chemical compounds and therefore have similar spectral signatures. The spectral variability of vegetation is mainly located within three to five wavelengths regions and the same number of well-placed wavebands could account for most of it. The usage of hyperspectral data to determine forest attributes is based on the assumption that improved identification of particular spectral features will lead to improved discrimination of forest attributes (Lefsky & Cohen, 2003). Curran (2001) states an idealised point of view of the advantage of a hyperspectral over a multispectral sensor in vegetation studies: Whereas a multispectral

Figure 1.3: The concept of imaging spectrometry (after Vane & Goetz, 1988).
sensor is suitable to identify the type of vegetation observed, the information about the condition of that vegetation resides within narrow spectral features which can be identified solely by sensors with many narrow wavebands.

Useful reviews of the application of imaging spectrometry are provided by Vane & Goetz (1993), Curran (1994), Hill & Megier (1994), Treitz & Howarth (1999) and Van der Meer & De Jong (2001). Past and current research in the field of ecology is mainly concentrating on estimating foliar chemical concentrations of plant leaves (Card et al., 1988; Elvidge, 1990; Curran et al. 1992; Jacquemoud et al., 1996; Fourty & Baret, 1998; Curran et al. 2001) and of canopies (Peterson et al., 1988; Wessman et al., 1988; Serrano et al., 2002; Zarco-Tejada et al., 2004), the estimation of structural canopy attributes (Gong et al., 2003; Lee et al., 2004), and cover type (Franklin, 1994; Martin et al., 1998). Despite intensive research and increasing literature available in the field of hyperspectral remote sensing, it is not yet clear, to what extent hyperspectral remote sensing a) can improve derivation of structural forest attributes compared to traditional sensor types and b) allows determination of chemical forest attributes.

1.3 Forest characteristics for ecosystem modeling and inventory

Terrestrial ecosystems play an important role in the exchange of matter and energy between soil, vegetation and the atmosphere (Mooney et al., 1987; Schimel et al., 2001). In this context, the Earth’s forests and woodlands are of particular importance, as they cover approximately 40 percent of the global land surface and provide around 70 percent of the annual net accumulation of terrestrial carbon (Peterson & Running, 1989). Most of the forests located in the mid-latitudes of the Northern Hemisphere act as a carbon sink, however, with considerable spatial and temporal variations and uncertainties (Valentini et al., 2000; Schimel et al., 2000; Canadell et al., 2000). In recent times, these forest ecosystems have been exposed to increasing pressure from environmental changes such as global warming and human population growth. With an increasing need to understand the role of forests in the global carbon budget (e.g. Kyoto protocol) and to assess the impact of environmental changes, these ecosystems have to be quantitatively analysed in order to better understand the prevailing processes. The most important processes occurring in forests include carbon exchange processes (photosynthesis and respiration), evapotranspiration, and nutrient cycling (Waring & Running, 1998). Already at the local level, these processes are difficult to measure (Baldocchi et al. 1996). At regional to global scales, processes are no longer measurable at all (Tenhunen et al., 1998). Consequently, models have been developed which simulate the fluxes of matter and energy. These simulation models require as inputs or for validation purposes accurate estimates of forest structural and chemical properties. Collecting this information by traditional field sampling methods is impracticable to do over regional to global scales (Wessman, 1994; Schimel, 1995). Using remotely sensed information about the
structural and chemical state of the canopy, it is possible to apply ecosystem simulation models to landscapes and regions (Ustin et al., 1991; Wicks & Curran, 2003). Canopy attributes relevant to ecosystem simulation modeling that can be derived from remote sensing are land cover, leaf area index (LAI), and standing biomass (Curran, 1994a). More recently, leaf nitrogen concentration (LNC) has been identified as another major state variable potentially open to remote sensing (Lucas & Curran, 1999). The accurate mapping of the forest into distinct classes of forest types is essential to parameterise an ecosystem model with a suitable set of parameters corresponding to each class (Hall et al., 1995).

One example of process-level ecosystem simulation model is Forest-BGC (Running & Coughlan, 1988). Forest-BGC (Forest-biogeochemical cycles) simulates the cycling of carbon, water, and nitrogen through forest ecosystems. Its generalised clone Biome-BGC (Kimball et al., 1997) is applicable for a range of ecosystems. To run the model requires the input of driving variables, state variables, and parameters. Climatic variables, such as temperature and precipitation are used to drive the model. State variables, such as LAI or transpiration, are controlled by the driving variables. LAI is used within the model to control the absorption of photosynthetically active radiation (PAR), estimate the canopy conductance and regulate the interception of precipitation (Waring & Running, 1999). LNC or leaf chlorophyll concentration controls the photosynthetic capacity of the foliage and, together with absorbed PAR, determines the upper limit for gross photosynthesis. The annual turnover of leaves is also correlated with LAI, LNC, and specific leaf area (Waring & Running, 1999). Finally the model uses parameters, such as specific leaf area or stomatal conductance that are species-specific and assumed to remain constant over time (Running & Gower, 1991).

Forest inventories provide relevant information about the composition, structure and productivity of forest stands (Fournier et al., 2003). Forest inventory data typically exist as quantitative attributes collected from ground surveys or as forest cover maps interpreted directly from aerial photographs. Ground survey attributes include tree heights, stem densities, volumes, species, age classes, among others. Traditionally, forest inventories have been repeated every 10 years or longer while forest ecosystem processes respond on shorter time scales (minutes, days, and years). To monitor the effects of these processes, observation strategies providing appropriate temporal resolutions are required (Cihlar et al., 2003).

In Germany, forest inventories typically exist for forest owned by the state or municipalities, but not for privately owned forest. For instance, in the state of Rhineland-Palatinate about a quarter of the total forest is private (Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft, 2004). As private forests may be of great economical and ecological importance, there is a considerable interest of the local and state forest authorities in such inventory data. Consequently, a remote sensing approach capable of mapping important forest attributes such as species type, age class, crown volume and biomass would be of great value.
1.4 Objectives and structure of thesis

The main objective of the research was to determine how forest attributes relevant to forest ecosystem simulation modeling and forest inventory purposes can be estimated from hyperspectral reflectance data. The study was restricted to forest plantations where stands are composed of trees of single species and age type to avoid problems related to mixed stands, stand heterogeneity and temporal dynamics. The investigated species, Norway spruce (*Picea abies* L. Kars.), is widespread throughout Europe and covers more than 50 percent of the Idarwald test site and is also the predominant tree species at Gerolstein test site.

The investigated attributes belong to three categories: Cover type, structural attributes and chemical attributes. The specific research objectives were:

- To investigate whether the forest cover type can be reliably derived from HyMap data using image classification. In this context data with broad to fine spectral resolution and large to small spatial resolution is compared (Chapter 5).

- To investigate the relationship between HyMap reflectance and structural canopy attributes with two different approaches (Chapter 6): 1) predictive models based on vegetation indices, 2) inversion of a forest reflectance model.

- To investigate the relationship between reflectance and chemical attributes using multivariate statistical analysis and to assess the reliability of scaling-up from leaf to canopy level (Chapter 7).

The outline of the study is as follows: Chapter 2 briefly describes the basic principles of optical remote sensing and the spectral properties of vegetation components followed by a review of the literature relevant to the remote sensing of forest attributes. Chapter 3 introduces the Idarwald and Gerolstein study sites and provides a description of the used material including field and laboratory measurements along with the acquisition and pre-processing of the image data. Chapter 4 outlines the methodologies used to derive the forest attributes from the reflectance data. Details are given on the employed image processing techniques, the empirical-statistical methods and the reflectance models. Chapters 5, 6, and 7 present and discuss the results obtained for cover type, structural attributes and chemical attributes. The different results are compared with respect to the applied methods and the characteristics of the input data. Chapter 8 provides conclusions derived from the research and closes with recommendations for future investigations. Summaries in English and German are given at the end of the work.
Chapter 2 Fundamentals

The aim of this chapter is to describe the fundamentals and review the literature relevant to the research. In the first section the basic principles of optical remote sensing are presented in brief. Atmospheric effects and their correction are discussed as it is required for a quantitative interpretation of airborne or satellite image data. Geometric effects are also considered. The next section describes the spectral properties of vegetation components with particular emphasis given to conifers at the needle, branch, and canopy level. In the last section the literature relevant to the remote sensing of forest attributes is reviewed.

2.1 Remote sensing

2.1.1 Physical principles

Remote sensing is defined as the extraction of information about an object without coming into physical contact with it (Schott, 1997). In the environmental sciences, remote sensing usually refers to the use of sensors to record images of the environment which can be interpreted to get useful information (Curran, 1985). This information is transferred from the object to the sensor through electromagnetic radiation. In the domain of passive optical remote sensing the Sun is the source of radiation. Radiation from the Sun reaches objects on the Earth’s surface and is reflected. The amount of reflected radiation depends on the properties of those objects. The reflected energy passing through the atmosphere is modified by scattering and absorption processes before it is recorded by a sensor. To understand how reflectance can be measured in remote sensing, basic physical and radiometric principles have to be considered. These are extracted from the textbooks of Curran (1985), Kraus & Schneider (1988), Schott (1997), and Lillesand & Kiefer (2000).

Wave theory describes electromagnetic energy as waves of wavelength $\lambda$ [μm] and frequency $\nu$ [s$^{-1}$] travelling at the speed of light $c$ [ms$^{-1}$].

$$c = \lambda \nu$$  \hspace{1cm} (2.1)

where wavelength is the distance between one peak and the next, frequency is the number of cycles that pass a fixed point in 1 s, and the speed of light in vacuum has a constant value of $2.998 \times 10^8$ ms$^{-1}$. It follows that waves with a short wavelength have a high frequency.
Quantum theory describes electromagnetic energy as being transferred in terms of energy packets or quanta. The energy $Q$ [J] of a quantum is given as

$$Q = h \nu = \frac{hc}{\lambda}$$  \hspace{1cm} 2.2

where, again, $\nu$ [s\(^{-1}\)] is the frequency and $h$ [Js] is the Planck’s constant with a value of 6.626\(\times\)10\(^{-34}\) Js. It can be concluded that waves with a short wavelengths have a high energy.

To compare measurements of energy over different periods of time, the radiant flux $\Phi$ [W] defines the rate at which the energy is propagating as the first derivative of the radiant energy $Q$ [J] with respect to time $t$ [s]:

$$\Phi = \frac{dQ}{dt}$$  \hspace{1cm} 2.3

The total energy flux radiated onto a unit area, e.g. of the Earth’s surface, from all directions is expressed as the irradiance $E$ [Wm\(^{-2}\)] and in the same way the total energy flux radiated away from a unit area in all directions is defined as the radiant exitance $M$ [Wm\(^{-2}\)]:

$$E = \frac{d\Phi}{dA}$$  \hspace{1cm} 2.4

$$M = \frac{d\Phi}{dA}$$  \hspace{1cm} 2.5

Both irradiance and exitance do not provide any angular or directional information of the radiant flux. To quantify what is measured by a sensor, the radiant flux $\Phi$ [W] is related to the unit projected area $dA \cos \theta$ [m\(^{-2}\)] and the unit solid angle $d\Omega$ [sr\(^{-1}\)]. The radiance $L$ [Wm\(^{-2}\)sr\(^{-1}\)] is defined as

$$L = \frac{d^2\Phi}{dA \cos \theta \cdot d\Omega}$$  \hspace{1cm} 2.6

As the flux is spectrally variable, each of the radiometric terms will vary with wavelength. Therefore, $L(\lambda)$ [Wm\(^{-2}\)sr\(^{-1}\)\(\mu m^{-1}\)] is the spectral radiance, $E(\lambda)$ [Wm\(^{-2}\)\(\mu m^{-1}\)] is the spectral irradiance, and $M(\lambda)$ [Wm\(^{-2}\)\(\mu m^{-1}\)] is the spectral exitance within a wavelength band.

To characterise an object in remote sensing, the spectral radiant exitance leaving the surface of the object $M_{\rho}(\lambda)$ has to be described as proportion of the incoming spectral irradiance $E(\lambda)$. This ratio, the spectral reflectance $\rho(\lambda)$, is the ability of the material to turn an incident flux back into the hemisphere above the material.

$$\rho(\lambda) = \frac{M_{\rho}(\lambda)}{E(\lambda)}$$  \hspace{1cm} 2.7
A plot of the spectral reflectance of an object as a function of wavelength is called a spectral reflectance curve. From the spectral characteristics described by the spectral reflectance curve conclusions can be drawn towards various qualitative and quantitative object properties. The wavelengths of reflected electromagnetic radiation particularly valuable in environmental remote sensing are in the visible, near infrared and mid infrared wavebands.

Similarly to reflectance, transmittance $\tau$ is the ratio of the transmitted exitance $M_\tau$ at the back of a sample to the incoming irradiance, and absorbance $\alpha$ is the ratio of the absorbed flux per unit area incident on a surface $M_\alpha$ to the irradiance onto the surface. Since the conservation of energy requires all the incident flux to be either reflected, transmitted, or absorbed, we have

$$\rho(\lambda) + \tau(\lambda) + \alpha(\lambda) = 1$$

Reflectance specifies the proportion of the incident energy which is reflected, but not the direction of the reflected radiation. Reflection can be either of the specular or diffuse type, or a mixture of both. In specular reflection, produced by smooth mirrorlike surfaces, all of the incident radiation is reflected towards the direction opposite to the incident angle. In that case, no interaction of the radiation with the surface takes place and little information is delivered to the sensor. The radiance reflected from diffuse reflectors, also called Lambertian surfaces, is the same in all directions. Most of the natural surfaces appear rough at visible and near infrared wavelengths and have a reflectance behaviour somewhat between the two extremes. Consequently, the reflected radiation reaching a sensor depends on the illumination and viewing geometry (section 2.2.4). However, Lambertian assumptions are often used as an approximation of reality. In that case, the relationship between the exitance $M$ from a surface and the radiance $L$ from that surface can be simplified, so that $L$ becomes independent of the observation angle $\theta$:

$$M = \int dM = \int L \cos \theta \, d\Omega = \pi L$$

Substitution of equation 2.9 into Equation 2.7 gives the spectral reflectance of Lambertian surfaces:

$$\rho(\lambda) = \frac{L(\lambda)}{E(\lambda) \pi}$$

There are basically two types of reflectance measurements, hemispherical and directional. In bi-hemispherical measurements the incident radiation is arriving from all directions and the reflected radiation is collected over all directions (laboratory measurement using an integrating sphere). In bi-directional measurements the angles of incidence and collection refer to one direction only (field or satellite measurement). As reflectance varies for non-Lambertian surfaces with illumination and viewing angle, a complete description of an object’s reflectance properties requires measuring the bi-directional reflectance at all possible
combinations of angles. This is achieved by the bi-directional reflectance distribution function (BRDF). To approximate a measurement of bi-directional reflectance in the laboratory or field, the ratio between the spectral radiance \( L(\lambda) \) at a certain angle to the object of interest and a diffuse reflector (with known reflectance properties) at the same angle is used instead.

Remote sensors, such as imaging spectrometers, record radiance values spatially registered and in a digital form. To extract spectral reflectance from remotely sensed images, corrections for the effects of the atmosphere and the image geometry have to be made.

### 2.1.2 Atmospheric effects

When radiation is passing through the atmosphere, it is partly absorbed or scattered. The element of fractional flux attenuated by extinction (absorption and scattering) at a certain location can be expressed as (Schott, 1997):

\[
\frac{d\Phi}{\Phi} = -\beta_{\text{ext}}(z, \lambda)dz
\]

2.11

where \( \beta_{\text{ext}}(z, \lambda) \) [m\(^{-1}\)] is the spectral extinction coefficient and \( dz \) [m] is an element of path length. For propagation along a finite path starting at distance zero where we have initial flux \( \Phi_0 \) to distance \( z \) where we have flux \( \Phi_z \), we have

\[
\int_{\Phi_0}^{\Phi_z} \frac{d\Phi}{\Phi} = \int_0^z -\beta_{\text{ext}}(z, \lambda)dz
\]

2.12

This can be solved for the simplified case of a homogenous medium to

\[
\tau = \frac{\Phi_z}{\Phi_0} = e^{-\beta_{\text{ext}}(\lambda)z}
\]

2.13

which is known as Lambert-Beer law. The term \( \Phi_z/\Phi_0 \) is a definition of transmittance \( \tau \). The product of the spectral extinction coefficient and path length \( \beta_{\text{ext}}(\lambda)z \) is generally referred to as the spectral optical depth or thickness \( \delta(\lambda) \) of a gas layer (Schott, 1997). It stands for the total extinction of a light beam passing through an air mass of thickness \( z \). Spectral extinction coefficient and spectral optical depth both depend on the density and size distribution of molecules and aerosols in the atmosphere. The spectral extinction coefficient \( \beta_{\text{ext}}(\lambda) \) is the sum of the absorption coefficient \( \beta_a(\lambda) \) and the scattering coefficients for Rayleigh scattering \( \beta_r(\lambda) \), Mie scattering \( \beta_a(\lambda) \) and non-selective scattering \( \beta_{\text{ns}}(\lambda) \):

\[
\beta_{\text{ext}}(\lambda) = \beta_a(\lambda) + \beta_r(\lambda) + \beta_a(\lambda) + \beta_{\text{ns}}(\lambda)
\]

2.14

Atmospheric absorption results in the loss of energy to atmospheric constituents. The most efficient absorbers of solar radiation are water vapour (H\(_2\)O), carbon dioxide (CO\(_2\)), and ozone (O\(_3\)). As these gases absorb radiation in specific wavelength bands, optical remote
sensing is limited to wavelength ranges in which the atmosphere is transmissive, the so-called atmospheric windows. Atmospheric scattering merely alters the direction of the radiation. Three types of scattering mechanisms can be distinguished: Rayleigh scatter occurs when radiation interacts with atmospheric molecules (N₂, O₂, Ar) that are much smaller in diameter than the wavelength of the interacting radiation. The Rayleigh scattering coefficient \( \beta_r \) is inversely proportional to the fourth power of wavelength \( \lambda \) (Kraus & Schneider, 1988):

\[
\beta_r(\theta, \lambda) = k \left( \frac{1}{\lambda^4} (1 + \cos^2 \theta) \right)
\]

where \( k \) is a constant factor and \( \theta \) is the scattering angle (deflection angle from beam direction). As short blue wavelength of about 400 nm are much stronger scattered than longer red wavelengths, the sky, which would appear black in the absence of scattering, appears blue. Mie scatter takes place when the atmospheric particle diameters are approximately equal to the wavelength of the radiation. This scattering effect, caused by aerosols such as water vapour and dust, depends on wavelength according to (Kraus & Schneider, 1988):

\[
\beta_m(\theta, \lambda) = k \lambda^{-b}
\]

where \( k \) is a constant factor and \( b \) varies between 0 and 4. As this function decreases slower

Figure 2.1: Interactions of incoming solar radiation with the atmosphere and the contamination of the satellite-measured radiance (from Tanré et al., 1985; modified after Hill, 1993). Refer to text for further explanations.
with increasing wavelength than the Rayleigh scatter function, Mie scatter tends to influence longer wavelengths. Non-selective scatter occurs when the atmospheric particles are much larger than the radiation wavelengths. It is caused by water droplets in clouds and fog which scatter all visible and near to mid infrared wavelengths about equally \( \beta_{ns}(\theta, \lambda) = \text{constant} \), resulting in a white appearance of clouds.

As the incoming solar irradiance interacts with the atmosphere due to the mentioned absorption and scattering processes, the signal measured at the satellite is contaminated. In order to correct a remotely sensed image, several atmospheric effects have to be considered. There are basically three relevant irradiance and three radiance terms (Figure 2.1): The direct irradiance term (1) is the part of the solar irradiance not affected by absorption or scattering arriving at the Earth’s surface. The scattered solar irradiance, the skylight (2), additionally illuminates the target. To a smaller extent, a fraction of already ground-reflected radiation that is again reflected from the atmosphere, arrive at the surface and is named environmental illumination (3). Together, these terms comprise the total irradiance onto the target. The satellite measured radiance \( L^* \) is mainly composed of the radiance reflected by the target \( L_{tar} \) (4) which is the part of the signal carrying the target information (target radiance). Additionally, a part of the incoming solar irradiance is backscattered towards the sensor without being in contact with the target before (atmospheric radiance, \( L_{atm} \)) (5). Finally, reflected irradiance from the environment surrounding the target \( L_{env} \) (background radiance) (6) is also directed into the sensor’s field of view (Tanré et al., 1985). Setting the at-satellite radiance to \( L^* = L_{tar} + L_{atm} + L_{env} \) and the irradiance onto the target to \( E = E_0 \cos \theta \), the at-satellite reflectance \( \rho^* \) is derived from equation 2.10 (ignoring atmospheric effects) as:

\[
\rho^* (\lambda) = \frac{\pi \cdot d_s^2 \cdot (L_{tar}(\lambda) + L_{atm}(\lambda) + L_{env}(\lambda))}{E_0(\lambda) \cos \theta} \tag{2.17}
\]

where \( E_0 \) is the extraterrestrial irradiance, \( \theta \) is the solar zenith angle and \( d_s \) is the earth-sun distance in astronomical units. Various researches have attempted to model the effects of the atmosphere on radiation with the aim of removing those effects from the signal. They developed so called radiative transfer models. For instance, Tanré et al. (1990) developed a model called the 5S code (Simulation of Satellite Signal in the Solar Spectrum). The 5S code computes the solar radiation reflected by the Earth-atmosphere system at satellite altitude in various illumination and observation conditions. The following simplifying assumptions are made (Hill, 1993): The atmosphere is considered to be horizontally homogenous and variations occur only in the vertical direction. The Earth’s surface is assumed to have Lambertian reflectance properties; bi-directional effects are neglected. In analogy to the abovementioned terms of radiance \( (L_{tar}, L_{env}, L_{atm}) \), the corresponding reflectance terms are used for convenience. The at-satellite reflectance \( \rho^* \) is expressed as a function of the target’s bi-directional reflectance \( \rho_{tar} \), the mean reflectance of the surrounding area \( \rho_{env} \), and the
atmospheric reflectance $\rho_{atm}$ (Tanré et al., 1990) (the wavelength dependence has been omitted for the sake of clarity):

$$
\rho^* = \left( \frac{\tau_{\text{dir}, \text{dif}} \left( \rho_{\text{atm}} \tau_{\text{dir}} \rho_{\text{sph}} \rho_{\text{sph}} \right) + \rho_{\text{env}} \tau_{\text{dif}} \tau_{\text{env}}}{1 - \rho_{\text{env}} \rho_{\text{sph}}} + \rho_{\text{atm}} \right) \tau_{\alpha \downarrow} \tau_{\alpha \uparrow}
$$

2.18

where $\tau_{\alpha \downarrow}$ is the transmittance due to absorption of radiation for the double path from the Sun to the Earth’s surface (downward) and from the Earth’s surface to the sensor (upward), $\tau_{\text{dir}}$ is the upward direct transmittance due to scattering, $\tau_{\text{dif}}$ is the upward diffuse transmittance due to scattering (both depend on the observation zenith angle), and $\tau_{\text{dir}, \text{dif}}$ is the downward direct and diffuse transmittance due to scattering (depends on the sun zenith angle). The factor $\rho_{\text{sph}}$ denotes the spherical albedo of the atmosphere; the term $(1 - \rho_{\text{env}} \rho_{\text{sph}})^{-1}$ increases $\rho^*$ due to radiance reflected by the surrounding area which is then again reflected by the atmosphere towards the target at the Earth’s surface, the so-called adjacency effect (Kaufman, 1989).

Target reflectance $\rho_{\text{tar}}$ is derived from the measured at-satellite reflectance $\rho^*$ by solving equation 2.18 for $\rho_{\text{tar}}$. The transmittance terms $\tau_{\text{dir}, \text{dif}}$, $\tau_{\text{dir}}$, $\tau_{\text{dif}}$, and the reflectance terms $\rho_{\text{atm}}$ and $\rho_{\text{sph}}$ can be calculated from the optical thickness $\delta(\lambda)$ (equation 2.13). While the Rayleigh optical thickness $\delta_r$ is spatially and temporally invariant (at constant air pressure), the aerosol optical thickness $\delta_a$ strongly varies in space and time and thus, is the main unknown variable in the atmospheric effects on remotely sensed imagery. Its value can be either derived from sun photometer measurements, inversion of the radiative transfer equation or from the image data itself (scene-based). One scene-based approach (Ahern et al., 1977) makes use of the principle that clear water is almost completely black in the near-infrared region of the spectrum. The at-satellite radiance measured in the near-infrared is assumed to be solely caused by scattering effects and consequently, can be used to estimate the aerosol optical thickness at this wavelength. This can be achieved in an iterative process using equation 2.18; the appropriate value of $\delta_a(\lambda)$ is found when the apparent reflectance of the water surface computed by the model equals the measured image reflectance. The approach has been successfully applied on multispectral satellite data (Hill, 1993).

### 2.1.3 Topographic effects

Measured reflectance of ground objects is, apart of the object’s reflectance properties, also influenced by the slope and the aspect of the image terrain. Slopes facing away from the sun appear relatively dark, and sun facing slopes with the same type of cover appear relatively bright. If the surface behaves as a Lambertian reflector, this topographic effect can be corrected using a cosine function which calculates the cosine of the solar zenith angle relative to the local terrain. However, the approach tends to produce overcorrected images especially when the angle between the Sun and the Earth’s surface is larger than 55° (Hill et al., 1995).
This can be attributed to the fact that most surfaces are of a non-Lambertian type and that the cosine correction can be applied only to the direct part of the solar irradiance. Therefore, correction of the topographic effect should be directly integrated into atmospheric correction models as they can provide realistic estimates of the direct and diffuse radiation (Hill et al., 1995).

### 2.1.4 Geometric effects

Raw digital image data usually contain geometric distortions that stem from the image acquisition process. The sources of these distortions are related to variations in the altitude, attitude, and velocity of the sensor platform or are due to factors such as panoramic distortion, earth curvature, atmospheric refraction, relief displacement, or non-linearities in the sweep of the sensor instantaneous field of view (Lillesand & Kiefer, 2000). The aim of geometric correction or image rectification is to make the image conform to a map projection system that represents the surface of the Earth on a plane. Such a procedure is essential to ensure that the position of samples collected in the field can be identified correctly in the image.

Imaging spectrometry data measured by airborne scanning systems particularly suffer from geometric distortions due to variations of the flight path and also of the attitude (roll, pitch, and heading) of the aircraft. These systematic, non-linear distortions can not be compensated for by a traditional polynomial image transformation. Instead, their correction has to be performed on a pixel wise calculation. The parametric geocoding software PARGE (Schläpfer et al., 1998) exactly reconstructs the scanning geometry for each image pixel based on position, attitude, and terrain elevation data and then puts each image pixel to its correct position. The position and attitude of the aircraft has to be recorded during the overflight. A digital elevation model (DEM) of high quality is required. Apart from its geocoding capability, PARGE can also integrate the DEM data into the geometry of the scanner to perform a correction of the terrain induced illumination effects (section 2.1.3).

### 2.2 Spectral reflectance properties of forests

A successful use of quantitative remote sensing in forestry requires an understanding of the spectral properties of the involved vegetation components. The interaction of electromagnetic radiation with the components of a forest canopy has to be considered on the molecular, cellular, leaf, branch, and canopy level. This summary is based primarily on reviews by Goel (1988), Guyot (1990), Baret (1991), Howard (1991), and Atzberger (1998).
2.2.1 The molecular level

Absorption bands in the spectra of materials are generally caused by electronic and vibrational processes. Isolated atoms and ions have discrete energy states. Electronic processes occur when photons of a specific wavelength are absorbed to shift electrons from a lower to a higher energy state. (Clark, 1999). In plant tissue, electronic transitions occur around the magnesium atom of chlorophyll resulting in strong absorption bands in the visible part of the spectrum. Chlorophylls are magnesium-containing porphyrins (organic pigments characterized by the possession of a cyclic group of four linked nitrogen-containing rings), chemically related to hemoglobin (Oxford Reference Online, 2004). Chlorophyll a (and similarly, chlorophyll b) has strong absorption bands in the visible domain at 430 nm and 660 nm.

When atoms form molecules, additional energy states are available through vibrational processes along the molecular bonds. The frequency of vibration depends on the strength of the bonds and the masses of the atoms bound together. For a molecule with N atoms, there are \(3N-6\) normal modes of vibrations called fundamentals. The energy transitions of fundamentals occur between the ground state and the first level excited vibrational state. Additional vibrations can also take place at approximately multiples of the original fundamental frequency. They are called overtones or harmonics when they involve multiples of a single fundamental mode, and combinations when different fundamental modes are involved (Clark, 1999).

The energy transitions of overtones are between the ground state and the second or third level excited vibrational state. Because transitions to higher energy states are successively less likely to occur, each overtone is successively weaker in intensity (Weyer, 1985). As the transition to second or third level energy states requires roughly twice or three time more

Figure 2.2: Absorption features of organic materials and water (from Weyer, 1985).
energy than it is needed to reach a first level state, the absorption bands occur at twice and three times of the frequency (one-half and one-third of the wavelength) of the fundamental (Weyer, 1985). The most prominent overtone bands observed in the near-IR and mid-IR domains are related to O-H, C-H, and N-H groups (stretch, bend, deformation). Combination bands usually involve stretch combined with one or more bending modes. As many different combinations are possible the near-IR to mid-IR region is very complex (Weyer, 1985).

Absorption bands caused by molecular vibrations are much weaker than those related to electronic processes. In plant tissue, water and organic materials (cellulose, lignin, protein, sugar, starch, tannin, and waxes) show absorption features related to vibrational processes (Figure 2.2). *Cellulose* is a polysaccharide (straight-chain D-glucose polymer) found in the cell walls of plants and forms 30 to 50 percent of the dry weight of most plants. It has near-IR/mid-IR absorptions at 1220, 1480, 1780, 1930, 2100, 2280, 2340, and 2480 nm (Elvidge, 1990). *Lignin* is a complex polymer of phenylpropanoid units and forms 10 to 35 percent of the dry weight in plant materials. It has near-IR/mid-IR absorptions at 1450, 1680, 1930 nm, and around 2100, 2300, and 2500 nm (Elvidge, 1990). *Starch*, a polysaccharide (D-glucose), is the principal food storage molecule in plants and not a major component of most leaf material. It has near-IR/mid-IR absorptions similar to those of cellulose plus absorptions at 990, 1560, and 1700 (Elvidge, 1990). *Proteins* contain many amino acids linked together by peptide bonds. The amide group (−CONH₂), formed by the reaction between adjacent carboxyl (−COOH) and amino (−NH₂) groups, contains C–N, C=O, and N–H bonds that all contribute to the characteristic absorptions related to the peptide linkage (Kokaly, 2001). The protein D-ribulose 1-5-diphosphate carboxylase is the most abundant nitrogen bearing compound in green leaves (30 to 50 percent of leaf nitrogen) and plays the central role in the carbon fixation in photosynthesis. It has minor absorptions in the mid-IR at 1500, 1680, 1740, 1940, 2050, 2170, 2290, and 2470 nm (Elvidge, 1990). Only leaf water has major absorption bands in the mid-IR and largely influences the overall reflectance in that spectral region. Leaf water can obscure absorption features related to protein (and other leaf chemicals) in green leaf materials with considerable consequences for its estimation from reflectance spectra. As protein is difficult to measure directly it is usually derived from total organic nitrogen by multiplying with a factor of 6.25 as plant protein roughly contain 16 percent nitrogen (Allen, 1991, cited from Kupiec, 1994). Absorption features related to nitrogen are summarised in Table 2.1.
2.2.2 The leaf level

Leaves represent the main surfaces of plant canopies where energy and gas are exchanged. Hence, their optical properties are essential to understand the transport of photons within vegetation (Despan & Jacquemoud, 2004). The general shape of reflectance and transmittance curves for green leaves is similar for all species (Figure 2.3). It is controlled by absorption features of specific molecules (section 2.2.1) and the cellular structure of the leaf tissue (Ustin et al, 1999).

The internal tissue of a leaf blade (lamina) is called mesophyll and consists of parenchyma cells. There are two distinct forms. *Palisade mesophyll* lies just beneath the upper epidermis.
and consists of cells elongated at right angles to the leaf surface. They contain a large number of chloroplasts and their principal function is photosynthesis. Spongy mesophyll occupies most of the remainder of the lamina. It consists of spherical loosely arranged cells containing fewer chloroplasts than the palisade mesophyll. Between these cells are air spaces leading to the stomata (Oxford Reference Online, 2005). Due to the random orientation of cells and the large number of air-cell interfaces, the diffusion of light mainly takes place in the spongy mesophyll (Guyot, 1990; Figure 2.4).

Three spectral domains can be distinguished: In the visible domain (400-700 nm) absorption by leaf pigments is the most important process leading to low reflectance and transmittance values. The main light absorbing pigments are chlorophyll a and b, carotenoids, xanthophylls, and polyphenols. Chlorophyll a is the major pigment of higher plants and together with chlorophyll b account for 65 percent of the total pigments. Chlorophyll a and b have absorption bands in the blue at around 430/450 nm and in the red domain at around 660/640 nm (Figure 2.5). These strong absorption bands induce a reflectance peak in the green domain at about 550 nm. Carotenoids and xanthophylls absorb mainly in the blue and are responsible for the colour of flowers, fruits, and the yellow colour of leaves in autumn. Polyphenols (brown pigments) absorb with decreasing intensity from the blue to the red and appear when the leaf is dead (Verdebout et al, 1994).

In the near-infrared domain (near-IR: 700-1300 nm) leaf pigments and cellulose are almost transparent, so that absorption is very low and reflectance and transmittance reach their maximum values. The level of reflectance on the near-IR plateau increases with increasing
number of intercell spaces, cell layers, and cell size. Scattering occurs mainly due to multiple refractions and reflections at the boundary between hydrated cellular walls and air spaces (Guyot, 1990).

The *red edge* is located between the absorption minimum in the red domain and the reflectance plateau in the near-IR region. This abrupt increase in reflectance, defined by the point of maximum slope, is generally referred to as the red edge (Horler et al., 1983). The wavelength position of the red edge, named *red edge inflection point* (REIP), occurs between 680 and 740 nm. The REIP depends on the amount of chlorophyll seen by the sensor (Dawson & Curran, 1998). The chlorophyll amount present in a vegetation canopy is characterised by the chlorophyll concentration of the leaves and the leaf area index (LAI). On a leaf level, solely the chlorophyll concentration determines the amount of chlorophyll. An increase in chlorophyll amount increases the chlorophyll absorption and broadens the associated chlorophyll-a absorption feature located at about 680 nm (Salisbury et al., 1987). As a consequence, the point of maximum slope (REIP) is moved towards longer wavelengths and a so-called red-shift of the inflection point is observed. Consequently, a decrease in chlorophyll amount will lead to a decrease of the chlorophyll absorption and will shift the REIP towards shorter wavelengths (blue shift).

In the *mid-infrared domain* (mid-IR: 1300-2500 nm), also called shortwave-infrared (SWIR), leaf optical properties are mainly affected by water and other foliar constituents. The major water absorption bands occur at 1450, 1940, and 2700 nm and secondary features at 960, 1120, 1540, 1670, and 2200 nm (Ustin et al, 1999). Water largely influences the overall reflectance in the mid-IR domain but also has an indirect effect on the visible and near-IR reflectances. Protein, cellulose, lignin, and starch also influence leaf reflectance in the mid-IR.
However, the absorption peaks of those organic substances are rather weak as they result from overtones or combinations related to fundamental molecular absorptions in the region 5 to 8 μm (Curran, 1989; section 2.2.1). The molecular absorptions are associated with certain chemical bonds, such as C-H, N-H, C-O, and O-H. In fresh leaves, spectral features related to organic substances are masked by the leaf water, so that estimation of leaf constituents is difficult (Verdebout et al, 1994).

Indirect factors, such as plant type, leaf age, and leaf health also determine the leaf optical properties as they can affect the leaf structure and the concentrations of absorbers. The leaf structure varies for different plant types (genera and species). Gymnosperms are mainly characterised by needle-shaped leaves. Angiosperms can be divided into the monocotyledons and the dicotyledons. Leaves from dicotyledon species reflect and transmit greater amounts of radiation than leaves of monocotyledon species of the same thickness (Guyot, 1990). Within a certain species, differences in structure occur between sun leaves and shade leaves (Howard, 1991). Leaf age strongly affects the optical properties of conifer needles as the chlorophyll content increases from year to year. Leaves of deciduous trees or annual plants, apart from the juvenile stage and senescence, have practically constant optical properties throughout the major part of their life (Guyot, 1990). Leaf health may be affected by mineral deficiencies, drought, pests and diseases. Drought can reduce the leaf water content; however, in natural conditions, it is necessary to have an extremely severe water stress to affect the leaf optical properties. Nitrogen deficiency reduces the chlorophyll content, the number of cell layers and thus, changes the whole reflectance spectrum (Guyot, 1990). Chlorosis, caused by conditions that prevent the formation of chlorophyll (e.g. lack of light or a deficiency of iron or magnesium) mainly affects leaf reflectance in the visible domain (Oxford Reference Online, 2005a). Plants affected by chlorosis are typically pale green or even yellow.
Leave reflectance also exhibits directional characteristics. For incidence angles normal to the leaf surface the reflectance is diffuse. For large incidence angles, reflectance is formed by a diffuse and a specular component; the latter is reflected by the cuticle and increases with increasing incidence angles. The specular component is practically the same in the visible and near-IR, whereas the diffuse component is low in the visible and high in the near-IR. This implies, that the relative portion of specular reflected radiation at the total reflected radiation is much lower in the near-IR than in the visible. Consequently, in the near-IR leaves can be considered Lambertian reflectors but in the visible they show strong directional effects (Guyot, 1990). Specularly reflected radiation will not contain any information on the internal biochemical composition of the leaf since no interaction has been occurred.

2.2.3 Soil optical properties

Apart from the leaf optical properties, soil optical properties also affect the reflectance of a forest canopy. In general, the reflectance of a bare soil increases progressively from the visible to the mid-IR and then decreases in the mid-IR with deep valleys at the water absorption bands. The main factors affecting soil reflectance are moisture content, soil texture, surface roughness, iron content, and organic matter content (Guyot, 1990).

With increasing soil moisture content the reflectance decreases for the whole spectrum due to the trapping effect. When particle size increases the soil reflectance increases for all the wavelengths, but the effect is weaker than that of moisture. Both soil moisture and soil texture are strongly related to each other: coarse, sandy soils are usually well drained, resulting in low moisture content and relatively high reflectances; fine, poorly drained clays typically have lower reflectances (Lillesand & Kiefer, 2000).

As a bare soil is not a Lambertian reflector, reflectance depends on the illumination and viewing geometry: reflectance values observed in the direction of the sun are enhanced (hot spot effect) whereas it is reduced in the direction opposite to the sun, because of the shadows.

![Figure 2.6: Classification of soil spectral curves into five general types (from Baumgardner et al., 1985).](image-url)
recorded. With increasing soil roughness reflectance decreases in the whole spectrum due to an enlarged occurrence of shadows (Baret, 1991).

Based on their spectral curve forms soils can be divided into 5 main types (Figure 2.6): organic dominated (a), minimally altered (b), iron affected (c), organic affected (d), iron dominated (e) (Baumgardner et al., 1985). Although dense forest canopies may mask the effect of the soil spectral curve on canopy reflectance spectra, factors such as fertility, drainage, and moisture-holding-capacity tend to influence the vegetation growth on a soil (Baumgardner et al., 1985) and thus, forest canopy reflectance.

2.2.4 The canopy level

The optical properties of a forest canopy depend mainly on the optical properties of the canopy constituents and on the canopy structure. The most important canopy elements are the leaves and the underlying soil. Optical properties of other parts of the plants, such as bark, flowers, fruit, etc. will not be considered in the following elaborations. When a plant canopy grows, the contribution of the soil to the observed total signal progressively decreases as the reflectance spectrum of the bare soil is gradually replaced by that of the plant. As a consequence, during growth the visible and middle-infrared (mid-IR) reflectances decrease and the near-infrared (near-IR) reflectances increase (Figure 2.7); the reverse phenomenon is observed during senescence or selective cutting of trees (Guyot, 1990). This behaviour is mainly due to the contrast between leaf and soil optical properties. In the red domain, soil

![Figure 2.7: Schematic representation of the changes in canopy reflectance during growth and senescence (after Guyot, 1990). See text for details.](image-url)
reflectance is generally higher than leaf reflectance whereas the reverse case is true in the near-IR domain (Baret, 1991). Under a dense forest canopy, the soil will be completely covered by tree crowns, understorey vegetation, and litter and is unlikely to contribute to the recorded signal. With a sparse tree canopy, the soil can be expected to contribute significantly (Howard, 1991).

The canopy structure is primarily defined by the leaf area index (LAI) and the leaf angle distribution (LAD). The LAI is defined as the total one-sided leaf area per unit ground area and represents the leaf amount of a canopy. When the amount of vegetation increases during growth, the canopy reflectance can reach saturation levels (Figure 2.8). In the case of annual crops, plots of reflectance as a function of the LAI show that in the visible the saturation level is reached for LAI of about 2 to 3, and in the near-IR for LAI around 5 to 6. These differences are mainly an effect of leaf transmittance, which is low in the visible and mid-IR while it is high in the near-IR (Guyot, 1990). In the near-IR, a number of leaf layers are required to obtain full extinction of the incident radiation while in the red domain almost no radiation penetrates through a single leaf layer (Baret, 1991). The value of LAI where saturation occurs also depends on the LAD: Plants with erect leaves need higher LAI values to saturate than plants having horizontal leaves (Guyot, 1990). For conifers, different observations were made. Peterson et al. (1987) found that the saturation level was reached at an LAI of approximately 8 in the red domain. Other results suggest a saturation of the NDVI at an LAI of about 5 (Chen & Cihlar, 1996) or no saturation effect at all in the case of the RVI (Chen et al., 2002). A possible explanation for these findings is that conifer stands with high LAI generally have

Figure 2.8: Schematic variation of canopy reflectance as a function of leaf area index for the red and near-IR domain (after Guyot, 1990; from Atzberger, 1998).
large tree crowns and large crown shadow fractions and thus appear dark in the red domain but bright in the near-IR domain due to strong multiple scattering, causing large RVI values (Chen et al., 2002).

LAD, the main factor characterizing the geometry of a plant canopy, is described by the leaf inclination and azimuth angles. The leaf inclination is the angle between the leaf normal (the perpendicular to the leaf surface) and the vertical axis. It generally changes from 0° for

Figure 2.9: Oblique (left) and nadir (right) view of black spruce (*Picea mariana*) demonstrating the gap effect. See text for details. From Sandmeier (2004). Photos by D. W. Deering.

Figure 2.10: Black spruce (*Picea mariana*) observed in backward scattering (left) and forward scattering (right) direction. See text for details. From Sandmeier (2004). Photos by D. W. Deering.
horizontal leaves (planophile) to 90° for vertical leaves (erectophile). The leaf azimuth angle is the orientation of the leaf towards a horizontal reference axis (e.g. north direction) and varies between 0° and 360°. The LAD changes the probability of gap through the entire canopy, which determines if the incident and outgoing radiation will be intercepted by the vegetation. LAD also determines the magnitude of the hot spot effect, which is the peak in the reflectance when the sun is directly behind the sensor. Another effect of canopy geometry occurs when leaves are clumped together rather than uniformly distributed. The clumping increases the probability of a gap which increases the influence of the vegetation in lower layers and of the soil. The reflectance of a forest canopy will also be influenced by the shape and size of the tree crowns (Goel, 1988).

As vegetation canopies are not perfect Lambertian reflectors the bi-directional reflectance of a canopy is affected by the illumination-viewing geometry. The illumination zenith angle is defined as the angle between the surface normal and the sun direction. It controls the amount of shadow present in a scene. The view zenith angle is defined as the angle between the surface normal and the sensor direction. It controls the relative portions of green vegetation and background seen by the sensor.

Bi-directional reflectance effects in vegetation canopies are primarily caused by the distribution of shadows. The amount of shadows observed under specific illumination and viewing angles is controlled the canopy architecture (LAI, LAD, and other structural

Figure 2.11: Causes for directional reflectance anisotropy of vegetation canopies and the associated typical shapes of the BRDF on the principal solar plane. From Lucht (2004).
attributes) which are influenced by plant species, phenology and health conditions. The influence of the canopy architecture on spectral BRDF data can be explained by the gap and backshadow effects and multiple scattering (Kimes, 1983; Sandmeier et al., 1998). The gap effect (Figure 2.9) occurs particularly in canopies with a vertically oriented (erectophile) structure such as grass lawns or conifer forests. At low view zenith angles, the lower less illuminated canopy levels are visible whereas at large view zenith angles, the well illuminated top layers of the canopy are seen. The gap effect alone leads to an azimuthally symmetric bowl shape centred around the nadir view angle (Sandmeier, 2004). The backshadow effect (Figure 2.10) causes the sensor to see the unilluminated shadowed vegetation surfaces in forward scattering direction. If only this effect was present in a scene, the lowest reflectance in a conifer forest would occur in the largest view zenith angle in forward scattering direction. In reality, the transmittance of leaves reduces the backshadow effect and under extreme viewing angles, the gap effect becomes dominant, shifting the minimum reflectance to small view zenith angles in forward scattering direction (Sandmeier, 2004). From the combination of gap and backshadow effect, densely vegetated erectophile canopies show i) the minimum reflectance near nadir view angles in the forward direction, ii) maximum reflectance (hot spot) in the backscatter direction at view zenith angle equal to the illumination zenith angle and iii) a general bowl shape of the BRDF (Figure 2.11). The bowl shape (a rise of reflectance with increasing view zenith angles) is caused by the gap effect and in dense forests, by mutual shadowing and obscuring of neighbouring trees (Meister et al. 2004).

The forward scattering component is increased by some specular reflection in most vegetation canopies. Both, gap and backshadow effects, are more pronounced at large illumination zenith angles and show stronger effects in spectral domains with lack of multiple scattering (e.g. red wavebands). The amount of multiple scattering regulates the darkness of the shadow in a vegetation canopy. In the near-IR and also the green, radiance reflected or transmitted by the leaves generates a diffuse secondary irradiance that brightens the shadow fractions and reduces the contrast between illuminated and shadowed fractions. Consequently, BRDF effects are less pronounced in the green and near-IR wavebands and strong in red and blue bands (Sandmeier, 2004).
Figure 2.12: Variation of simulated forest canopy reflectance as a function of: A) leaf chlorophyll a+b content \( \text{mg m}^{-2} \), B) leaf water content \( \text{g m}^{-2} \), C) nitrogen content \( \text{g m}^{-2} \), D) single tree leaf area index \( \text{m}^2\text{m}^{-2} \), E) average leaf inclination angle \( \circ \), F) stem density \( \text{ha}^{-1} \), G) tree height \( \text{m} \), H) crown diameter \( \text{m} \). Spectral reflectances were simulated by the INFORM model using the following parameter values: \( C_{AB} = 200 \text{ mg m}^{-2} \), \( C_W = 50 \text{ g m}^{-2} \), \( C_N = 1 \text{ g m}^{-2} \), \( \text{LAI}_s = 6 \text{ m}^2\text{m}^{-2} \), \( \text{ALA} = 55\circ \), \( \text{SD} = 700 \text{ ha}^{-1} \), \( H = 10 \text{ m} \), \( \text{CD} = 5 \text{ m} \), \( \theta_s = 35\circ \), \( \psi = 0\circ \), \( \text{sky} = 0.1 \).
Reflectance models are useful tools for analysing the effects of forest canopy attributes on reflectance signatures. By analysing the sensitivity of forest reflectance to the input variables, one can assess the error that may be caused by inaccurate input data. The variation of forest canopy reflectance with changes in certain canopy attributes as modelled by the INFORM model is shown in Figure 2.12.

2.2.5 Spectral reflectance properties of Norway spruce

Spectra of Norway spruce and other tree species show a dramatic decrease of reflectance from the needle to the branch to the canopy level in the near-IR and visible domain (Figure 2.13). In the visible changes in reflectance can be attributed to the influence of increased shadowing within the total scene from the needle to the canopy level. In the near-IR region, although reflectance generally increases with increasing canopy density or LAI due to the high reflectance and transmittance, four significant factors may account for a reduction in near-IR reflectance at the branch and canopy level: i) shadowing; ii) absorption of radiation by the twigs, branches, bark, and litter; iii) needle age; and iv) the clustering of needles on the branches which may enhance the trapping of radiation (Williams, 1991). With increasing LAI (or number of branches), reflectance in the near-IR domain is increasing; saturation occurs at an LAI of about 7-9 (Figure 2.14).

Figure 2.13: Measured reflectance spectra of Norway spruce at the needle, branch, and canopy level (modified from Williams, 1991).
2.2.6 Age dependence of forest reflectance

Forest age largely controls the structural parameters that determine the spectral reflectance of a forest stand. The age dependence of Norway spruce forest reflectance has been derived from airborne reflectance measurements and through forest reflectance model simulations by Nilson & Peterson (1994). The analysis of airborne data revealed a relatively fast reflectance decrease up to a saturation point at 20 years and 80 years for red and near-IR reflectances, respectively (Figure 2.15). This behaviour was explained by the fact that the amount of shade increases with stand age as the roughness of the forest canopy is increasing. Generally, simulated reflectances agreed with measured reflectances; however, some essential differences were observed by the authors: i) simulated reflectance slightly increased from an age of 30 onwards (red and near-IR domain), ii) in the near-IR domain, simulated reflectance increased during the first 10 years of stand development. Despite a certain disagreement between measured and simulated reflectances, simulated age courses of the scene component proportions (sunlit crowns, shaded crowns, sunlit ground, and shaded ground) largely explained the observed temporal reflectance variation: During the first 20-30 years of stand development, where a fast decrease in reflectance was observed (particularly in the red domain), the proportion of sunlit crowns increased (with increasing canopy cover), whereas the proportion of sunlit ground dramatically decreased. The proportion of shaded ground was constantly increasing with stand age and could be responsible for the observed reflectance decrease in the near-IR domain up to the end of the succession (Nilson & Peterson, 1994). However, the proportion of the shaded crowns did not increase between 40 and 100 years in
Nilson & Peterson (1994) identified the following driving factors of the reflectance age course of a forest stand from their simulation studies: species composition of the main tree storey, canopy cover, tree-storey LAI, understorey vegetation, leaf/supporting tissue ratio in the tree storey. According to the authors, also the height distribution of the tree layer (shade effects) and the needle age distribution (effect of different needle optical properties) contribute to the temporal reflectance changes.

2.3 Remote sensing of forest attributes

2.3.1 Classes of forest attributes

There are several classes of attributes to be remotely sensed in forests that share related mechanisms of, and limitations on, their remote estimation. Lefsky & Cohen (2003) defined classes of attributes related to: canopy cover, stand structure, stand composition, and disturbance. Attributes related to canopy cover include foliage or canopy cover, fraction of absorbed photosynthetically active radiation ($F_{\text{APAR}}$), and leaf area index (LAI). The attributes related to cover, although their formal definition and their estimation in the field differ, have a
similar, clear link with the remotely sensed signal: An increase of a cover related attribute will be indicated by decreased reflectance in the red and increased reflectance in the near-IR. However, the relationship between canopy cover and LAI is asymptotic; with each additional layer of leaves the amount of additional cover becomes smaller. As changes in reflectance are mostly driven by change in cover, the relationship between reflectance and LAI is also asymptotic (Lefsky & Cohen, 2003; Figure 2.8). Attributes related to cover have been most frequently the target of research in terrestrial remote sensing (section 2.3.3). Attributes related to canopy height or stand age include tree volume, aboveground biomass, basal area, diameter-at-breast-height, and stem density. With the exception of very low productivity stands, canopy closure occurs early in most stands while attributes related to canopy height continue to change. Estimation of these attributes is limited to the period of changing canopy cover. Thus, accurate estimation of height-related attributes is one of the most difficult tasks in the remote sensing of moderate to high biomass forests. In some conifer stands, the structural canopy properties that can be observed continue throughout stand development, and determination of height-related variables is more successful (Lefsky & Cohen, 2003). Attributes related to composition comprise physiognomy (e.g. trees, shrubs, and herbs), phenological types (deciduous and evergreen), species composition, and various cover type classifications (such as stand age classes). Attributes related to disturbance include attributes useful to monitor stand replacement, fires, thinning, etc. Within this research work the classification proposed by Lefsky & Cohen (2003) was modified. Compositional attributes were denoted as attributes related to forest type. Both, attributes related to cover and attributes related to canopy height, were combined to structural forest attributes. Additionally, a class denominated chemical forest attributes comprising concentrations or contents of leaf constituents was created. Attributes related to disturbance involve the analysis of multitemporal data sets and were not considered.

2.3.2 Attributes related to forest type

Multispectral satellite data have shown a good capability for mapping broad forest type classes, such as hardwoods, softwoods, mixed woods, open wetlands and clearcuts (Bryant et al., 1980). Some conifer species discrimination capabilities have been reported, but consistent discrimination has not been possible (Leckie, 1990). For instance, Walsh (1980) identified seven classes of coniferous tree species through the use of Landsat MSS data with an average accuracy of 88.8 percent. Results obtained by Coleman et al. (1990) revealed that TM imagery was inadequate for differentiating among pine species, but suitable to differentiate five pine age classes (four age classes out of the five age classes had a class mean of less than 20 years, overall classification accuracy was 59.2 percent and 44.6 percent for two sites). Separation of age classes in spruce stands younger than 50 years was achieved by Vohland (1997) using TM data, but discrimination of older stands proofed to be problematic.
Few studies have reported on the use of hyperspectral data in mapping forest species. Franklin (1994) compared panchromatic SPOT, multispectral TM and hyperspectral CASI data for discrimination of 9 types of forest species; the obtained overall classification accuracy were 49.2 percent using SPOT data and about 81 percent for both TM and CASI data. In attempting to discriminate age classes of Douglas fir with CASI data, Niemann (1995) concluded that a limited separation based on age is feasible but that discrimination of stands older than 40 years is not possible. Image classification of 11 forest cover types, including pure and mixed stands of deciduous and conifer species using AVIRIS data yielded an overall classification accuracy of 75 percent (Martin et al., 1998). The study used 11 spectral bands for species classification which were closely correlated with field measured canopy nitrogen and lignin concentrations. Prediction of forest age from TM data and additional topographic information using neural networks resulted in rms errors about 5 years for 50 year old stands (Kimes et al., 1996). In situ hyperspectral data has also been explored to discriminate conifer (Gong et al., 1997) and tropical species (Cochrane, 2000).

It can be concluded that detailed information suitable for forest inventories can only partly be derived from remote sensing data. While discrimination of young to medium aged forest stands was shown to be possible with both, multi- and hyperspectral data, separation of medium aged to old stands could not be achieved.

### 2.3.3 Structural attributes

#### Field estimation of leaf area index

To derive LAI maps from remote sensing data accurate field estimation of LAI is essential. Canopy LAI is difficult to assess in forests through direct methods, such as stratified clipping techniques, litter collections, and allometric relationships. These methods are often site specific and have a limited spatial and temporal validity (Soudani et al, 2002). For instance, allometric relationships between LAI and diameter at breast height are inherently unresponsive to seasonal change, as they are typically derived from variables collected at various times throughout the growing season (Lucas, 1995). To overcome such difficulties, indirect optical methods have been developed. Ground-based LAI measurements using optical instruments have become a standard as they allow for intensive spatial and temporal sampling over large areas (Soudani et al, 2002). The LAI-2000 plant canopy analyser (Li-Cor Inc., Lincoln, USA) (Welles & Norman, 1991) has an optical sensor that consists of five detectors, arranged in concentric rings, which measure radiation (below 490 nm) from different sections of the sky. Canopy transmittance for these different sections is calculated as the ratio between below-canopy and above-canopy readings for each detector ring (Stenberg, 1996). From these transmittance values (gap fraction) LAI is estimated using an approximation of the Lambert-Beer law (see section 2.1.2). This estimation is based on the assumption of randomly
distributed leaves in the crown and does not consider other shading elements such as stems. In reality, however, conifer stands show a non-random leaf distribution, as needles are grouped (clumped) first in shoots and shoots are further grouped into branches and crowns (Chen et al., 2002). Therefore, LAI estimates from optical methods are effective plant area indices (LAI_{eff}), which account neither for clumping nor for separating leaves from non-green material, nor for distinguishing species (Soudani et al., 2002). LAI_{eff} represents the equivalent leaf area of a canopy with a random foliage distribution to produce the same light interception as the true LAI (Fernandes et al., 2002). LAI_{eff} generally differs considerably from the true LAI. An underestimation of the true LAI results from clumping as it increases the canopy gap fraction for a given leaf area. An overestimation can be attributed to the fact that the instrument considers not only the leaves but also woody canopy parts such as stems (Stenberg, 1996). Several correction factors were developed to derive the true LAI from LAI_{eff} in forests (Fassnacht et al., 1994; Stenberg, 1996; Chen, 1996). According to Chen & Cihlar (1996), from a remote sensing perspective LAI_{eff} has several advantages over the true LAI: first, the effective LAI was better correlated to vegetation indices (RVI, NDVI) than the true LAI (Chen & Cihlar, 1996); second, the true LAI varied by about 10-30% while LAI_{eff} remained virtually unchanged in the course of a year as the new needles grew on top of the old needles and did not reduce the canopy gap fraction significantly (Chen, 1996). Therefore, in this research work the focus is on mapping LAI_{eff} instead of the true LAI. As suggested by Fernandes et al. (2002) it is possible to convert maps of LAI_{eff} to true LAI using correction factors estimated at the site or given in the literature. For simplicity, in this study the term LAI is used when referring to either effective or true LAI.

**Empirical approaches**

The majority of studies for extracting structural attributes from remotely sensed data have used empirical techniques to relate the spectral measurements to the variable of interest (Treitz & Howarth, 1999). Spectral measurements are usually converted to spectral vegetation indices (VIs) as these simple transformations are more closely related with plant biophysical properties. If structural parameters are strongly correlated with remotely sensed VIs, the remote sensing imagery can be used to predict those structural characteristics over large areas. A lot of research has been done on the estimation of forest LAI from remote sensing data within about the last 15 years. A selection of literature results is presented in Table 2.2. Most of the studies on forests have used broadband VIs to derive LAI of conifer forest stands; however, with varying success. Several authors found good correlations between the normalized difference vegetation index (NDVI) or the ratio vegetation index (RVI) derived from TM or AVHRR images and field-measured LAI (Running et al., 1986; Peterson et al., 1987; Spanner et al., 1990a; Curran et al., 1992a; White et al., 1997; Franklin et al., 1997). In all of these studies the range in canopy LAI was very large. Those studies where the LAI
range was rather limited generally show a smaller correlation or no relationship (Badhwar et
al., 1986; Spanner et al., 1990; Herwitz et al., 1990; Hall et al., 1995; Chen & Cihlar; 1996).
Better relationships where found even for small variations in canopy LAI when using
modified VIs, such as the reduced simple ratio or the normalised difference between near-IR
and mid-IR (Brown et al., 2000; Chen et al., 2002; Boyd et al.; 2000) or when multiple linear
regression was employed (Fassnacht et al., 1997).

Few studies looked at the suitability of high spectral resolution remote sensing data to derive
the LAI of conifer forests. Gong et al. (1995) tested CASI data using univariate and
multivariate regression and a VI based algorithm and found strong relationships with
reasonable low errors. Danson & Plummer (1995) tested if the red edge inflection point
(REIP) is primarily controlled by forest canopy LAI using helicopter-borne spectroradiometer
data and found a strong non-linear correlation between plot LAI and the REIP for Sitka spruce
(Picea sitchensis). For the same tree species, forest LAI was recently related to the canopy
REIP computed from imaging spectrometer data (CASI) with success (Lucas et al., 2000).

A systematic investigation on the performance of various broadband multispectral and narrow
band hyperspectral vegetation indices has been done on agricultural crops (Boegh et al., 2002;
Broge & Mortensen, 2002) and very recently, also in forests. For instance, Gong et al. (2003)
estimated forest LAI using vegetation indices derived from Hyperion hyperspectral data. Lee
et al. (2004) performed a comparative analysis of hyperspectral versus multispectral data for
estimating LAI in four different biomes. Using canonical correlation analysis, the authors
found that individual, narrow bands of AVIRIS data yielded better relationships with LAI
than broadband data for grassland and forest biomes. Given the few studies undertaken, it is
still not sufficiently investigated whether the high spectral resolution data offer advantages
over broadband data.

Generally, relationships between the reflectance of a vegetation canopy and its LAI can suffer
from effects of internal and external variables. For instance, Spanner et al. (1990; 1990a)
found linear and strongly positive relationships between near-IR radiance and LAI only for
stands with a canopy cover greater 89% whereas old growth stands with incomplete
overstories had low near-IR radiances as a result of less reflective forest litter and shadows.
However, in more open stands the exhibition of highly reflective understorey vegetation could
also increase near-IR radiance. Soudani et al. (2002) found that the understorey LAI
contributed 37-42% to the total LAI in pine forests. Chen & Cihlar (1996) observed that late
spring TM images were superior to summer images for determining overstorey LAI in boreal
conifer stands because the effect of the understorey was minimized in the spring before the
full growth of the understorey and moss cover. On semi-natural forest sites with a wide range
of green vegetation amount, canopy cover or LAI was considered the most important variable
to control canopy reflectance, even with variable understorey (e.g., Spanner et al., 1990a). On
the contrary, in coniferous forest plantations, which are managed to maintain a large amount of green vegetation with little spatial variation (and thus, low ranges in LAI) the relation between LAI and canopy reflectance may be weaker, as also other structural stand characteristics (stem density, tree height) determine the stand spectral response (Danson & Curran, 1993; Treitz & Howarth, 1999).

To overcome these problems various strategies were suggested. Through stratification by forest cover type, Franklin et al. (1997) yielded more precise LAI estimates from TM data. They also found that the GIS-based implementation of additional information on stem density

Table 2.2: Selected literature results on the estimation of forest structural attributes. Sensors: ATM = Airborne Thematic Mapper, TM = Thematic Mapper, AVHRR = Advanced Very High Resolution Radiometer, CASI = Compact Airborne Spectral Imager, DAIS = Digital Airborne Imaging Spectrometer; Vegetation Index (VI): RVI = Ratio Vegetation Index, NDVI = Normalised Difference Vegetation Index, RSR = Reduced Simple Ratio, REIP = Red edge inflection point, VI3 = (near-IR – mid-IR) / (near-IR + mid-IR); Method: BRA = Bivariate regression analysis, MLR = Multiple linear regression, IRM = Inversion of reflectance model, CCA = Canonical correlation analysis

<table>
<thead>
<tr>
<th>Sensor</th>
<th>VI/Method</th>
<th>Variable</th>
<th>Range</th>
<th>N</th>
<th>Relation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM</td>
<td>RVI/BRA</td>
<td>LAI (various conifers)</td>
<td>1-16</td>
<td>18</td>
<td>( R^2 = 0.82 )</td>
<td>Running et al., 1986, Peterson et al., 1987</td>
</tr>
<tr>
<td>TM</td>
<td>( \rho_{red}, \rho_{near-IR} /BRA )</td>
<td>LAI (aspen)</td>
<td>0.5-3.5</td>
<td>29</td>
<td>No relation</td>
<td>Badhwar et al., 1986</td>
</tr>
<tr>
<td>AVHRR</td>
<td>NDVI/BRA</td>
<td>LAI (various conifers)</td>
<td>1-13</td>
<td>19</td>
<td>( R^2 = 0.70-0.79 )</td>
<td>Spanner et al., 1990</td>
</tr>
<tr>
<td>TM</td>
<td>NDVI/BRA</td>
<td>LAI (various conifers)</td>
<td>1-16</td>
<td>73</td>
<td>( r = 0.60 )</td>
<td>Spanner et al., 1990a</td>
</tr>
<tr>
<td>TM</td>
<td>RVI/BRA</td>
<td>LAI (pine)</td>
<td>4-7</td>
<td>15</td>
<td>No relation</td>
<td>Herwitz et al., 1990</td>
</tr>
<tr>
<td>TM</td>
<td>NDVI/BRA</td>
<td>LAI (pine)</td>
<td>2-10</td>
<td>16</td>
<td>( R^2 = 0.35-0.86, ) ( rmse = 0.74 )</td>
<td>Curran et al., 1992</td>
</tr>
<tr>
<td>CASI</td>
<td>NDVI/MLR</td>
<td>LAI (various conifers)</td>
<td>1-12</td>
<td>30</td>
<td>( R^2 &gt; 0.80, ) ( rmse &lt; 1.0 )</td>
<td>Gong et al., 1995</td>
</tr>
<tr>
<td>Helicopter-borne data</td>
<td>REIP/BRA</td>
<td>LAI (spruce)</td>
<td>5-11</td>
<td>14</td>
<td>( r = 0.91 )</td>
<td>Danson &amp; Plummer, 1995</td>
</tr>
<tr>
<td>TM</td>
<td>NDVI/BRA</td>
<td>LAI Cover (various conif.)</td>
<td>0.5-3.5</td>
<td>22</td>
<td>( R^2 = 0.38-0.66 ) ( R^2 = 0.26-0.63 )</td>
<td>Chen &amp; Cihlar, 1996</td>
</tr>
<tr>
<td>TM</td>
<td>RVI/MLR</td>
<td>LAI (various conifers)</td>
<td>1.0-4.5</td>
<td>10</td>
<td>( R^2 = 0.71 )</td>
<td>Fassnacht et al., 1997</td>
</tr>
<tr>
<td>TM</td>
<td>RVI, NDVI /BRA</td>
<td>LAI (various vegetation types)</td>
<td>2-15</td>
<td>39</td>
<td>( R^2 = 0.65-0.90 )</td>
<td>White et al., 1997</td>
</tr>
<tr>
<td>TM</td>
<td>NDVI/BRA</td>
<td>LAI (various forest types)</td>
<td>Not specified</td>
<td>17</td>
<td>( R^2 = 0.15 ) (all spec.) ( R^2 = 0.93 ) (conifers)</td>
<td>Franklin et al., 1997</td>
</tr>
<tr>
<td>TM</td>
<td>IRM</td>
<td>LAI (various vegetation types)</td>
<td>0-7</td>
<td>Not specified</td>
<td>LAI pattern in good accordance to the land use map</td>
<td>Kuusk, 1998</td>
</tr>
</tbody>
</table>
increased the accuracy of the LAI estimates and that multiple linear regressions using NDVI and stem density improved correlations with LAI in comparison to the simple correlations. Chen et al. (2002) identified accurate information about subpixel mixture of the various cover types as a key to improving the accuracy of LAI estimates. Badhwar et al. (1986) recommended the use of canopy reflectance models to understand the understorey effects on forest reflectance.

### Physically based approaches

It is believed that simple empirical relationships need to be replaced by physically based forest reflectance models as a model inversion approach may provide a more robust method for estimating structural characteristics (Gemmell et al. 2002, Treitz & Howarth, 1999). The effects of internal and external variables might be taken into account in the inversion process if such information is available.

In particular in the case of forests, physically based approaches have not yet been widely tested. In the forward mode the CR model computes the spectral reflectance for a certain set of leaf or canopy parameters. Forward modeling provides an efficient way to investigate the effects of forest canopy variables on reflectance signatures and may improve the understanding of how the canopy structure influences the radiation field (e.g., Rautiainen et al., 2004; Schlerf & Atzberger, 2002). For the retrieval of those canopy attributes from measured signals it is necessary to invert the model. Inversion has been mostly considered as

### Table 2.2 (continued)

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Band</th>
<th>LAI (Species)</th>
<th>Range</th>
<th>N</th>
<th>$R^2$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM</td>
<td>RSR/BRA</td>
<td>LAI (pine, spruce)</td>
<td>0.5-6.0</td>
<td>35</td>
<td>$0.55-0.70$</td>
<td>Brown et al., 2000</td>
</tr>
<tr>
<td>AVHRR</td>
<td>NDVI/BRA</td>
<td>LAI (pine, aspen, spruce)</td>
<td>2.0-4.0</td>
<td>10</td>
<td>$0.46$</td>
<td>Boyd et al., 2000</td>
</tr>
<tr>
<td>CASI</td>
<td>REIP/BRA</td>
<td>LAI (spruce, pine)</td>
<td>3-12</td>
<td>16</td>
<td>$0.94$</td>
<td>Lucas et al., 2000</td>
</tr>
<tr>
<td>CASI</td>
<td>IRM</td>
<td>LAI (various conifers)</td>
<td>1.0-3.0</td>
<td>20</td>
<td>$0.51-0.86$</td>
<td>Hu et al., 2000</td>
</tr>
<tr>
<td>CASI</td>
<td>IRM</td>
<td>LAI (pine)</td>
<td>1.0-3.0</td>
<td>20</td>
<td>$0.16-0.67$</td>
<td>Fernandes et al., 2002</td>
</tr>
<tr>
<td>TM</td>
<td>IRM</td>
<td>Cover (pine)</td>
<td>0.1-1.0</td>
<td>33</td>
<td>$0.33-0.38$</td>
<td>Gemmell et al., 2002</td>
</tr>
<tr>
<td>AVHRR</td>
<td>RSR/BRA</td>
<td>LAI (various conifers)</td>
<td>0.5-10.5</td>
<td>&gt; 50</td>
<td>$0.62$</td>
<td>Chen et al., 2002</td>
</tr>
<tr>
<td>Hyperion</td>
<td>VI</td>
<td>LAI</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Not specified</td>
</tr>
<tr>
<td>AVIRIS</td>
<td>CCA</td>
<td>LAI (various biomes)</td>
<td>1.9-5.1</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Not specified</td>
</tr>
<tr>
<td>CASI</td>
<td>IRM</td>
<td>LAI (pine, spruce)</td>
<td>1-6</td>
<td>46</td>
<td>Relative rmse = 14%</td>
<td>Fernandes et al., 2004</td>
</tr>
<tr>
<td>DAIS</td>
<td>IRM</td>
<td>LAI (poplar)</td>
<td>1.0-3.5</td>
<td>26</td>
<td>rmse = 0.40</td>
<td>Meroni et al., 2004</td>
</tr>
<tr>
<td>DAIS</td>
<td>IRM</td>
<td>LAI (pine)</td>
<td>2-4</td>
<td>4</td>
<td>rmse = 0.5</td>
<td>Kötz et al., 2004</td>
</tr>
</tbody>
</table>
a way to validate radiative transfer models but has rarely been used as an alternative to extract canopy attributes from measured reflectance (Jacquemoud et al., 2000). Most of these works concern the retrieval of canopy structural attributes (LAI, leaf angle distribution) by using multi-angle reflectances (Goel & Thompson, 1984; Goel & Deering, 1985; Iaquinta et al., 1997; Bicheron & Leroy, 1999). Recently, hyperspectral reflectances have been increasingly used as an input to model inversion for estimation of structural and chemical vegetation attributes (Jacquemoud et al., 2000; Hu et al. 2000; Zarco-Tejada et al., 2001; Meroni et al., 2004; Fernandes et al., 2004; Kötz et al., 2004).

The majority of the aforementioned and other inversion studies focused on agricultural crops but only some on forests. According to Zarco-Tejada et al (2001), until recent, no validation has been found in the literature reporting results in forest canopies from airborne or satellite-measured reflectance with ground truth. For instance, Woodcock et al. (1994) estimated tree size and cover from TM data using the Li-Strahler model but results were not validated with measured field data. Kuusk (1998) inverted the MCRM model on reflectance in five TM bands to estimate LAI, average leaf angle, and chlorophyll content for various vegetation types. Ground truth data was not available. He found the estimated LAI to be in good accordance with the land use map but had to concede that the LAI particularly of conifer forests was probably underestimated by the model.

For the past five years, an increasing number of successful LAI and crown cover estimation results were reported. Gemmell & Varjo (1999) were unable to provide accurate estimates of LAI and canopy cover in boreal forests; when background reflectance was considered, estimation of canopy cover succeeded with good accuracy (Gemmell, 1999). The use of multi-angle ASAS data further improved estimates of forest canopy cover (Gemmell, 2000). Hu et al. (2000) derived LAI of boreal forest canopies through inversion of a modified FLIM model using hyperspectral CASI data. Validation results showed moderate to high correlations between estimated and field-measured canopy LAI; increased correlations were found when the understorey reflectance was allowed to vary within the range of variability as determined by field measurements. In combination with a multispectral clustering approach, the modified FLIM was used to map spatial patterns of effective LAI within boreal forest stands. Depending on the site, correlations between estimated and measured LAI were poor to moderate (Fernandes et al., 2002). Zarco-Tejada et al. (2001) found acceptable relations between estimated and measured leaf pigment content of closed forest canopies using CASI data and coupled SAILH and PROSPECT models. Meroni et al. (2004) estimated LAI of poplar plantations from DAIS data using the SAIL model coupled with the PROSPECT model and obtained an rmse of 0.40-0.45 against ground truth. Fernandes et al. (2004) report a relative rmse of 14% for black spruce sites using CASI data and the FLIM-CLUS algorithm. Kötz et al. (2004) report an rmse of 0.5 for estimating LAI in pine stands from DAIS data through inversion of a coupled GeoSAIL and PROSPECT models.
Due to the complex character of a CR model an analytical solution is not possible and inversion strategies had to be developed. A successful inversion of a reflectance model requires three factors: a good model, an appropriate inversion procedure, and a set of calibrated reflectances (Jacquemoud et al., 2000). Traditional inversion of reflectance models employs an optimisation technique to estimate the model parameters by minimising a merit function (Goel, 1989). An iterative process is necessary to find the optimal estimates of these parameters. The main drawbacks of this method are i) difficulty in achieving globally optimal and stable results, ii) difficulty in retrieving more than 2 parameters simultaneously (Gong et al., 1999), and iii) computational inefficiencies that prohibit an operational application on a per-pixel basis for regional or global studies (Kimes et al., 2002).

Recently, Artificial Neural Networks (ANN) have been employed for reflectance model inversion (Gong et al., 1999; Udelhoven et al., 2000; Kimes et al., 2002) in order to overcome the above mentioned limitations. Gong et al. (1999) found the Coupled Atmosphere and Canopy (CAC) model extremely difficult to use for the retrieval of structural parameters with ordinary (numerical or look-up table) inversion methods. They employed an ANN on simulated multi-angle reflectances and were able to retrieve many parameters simultaneously; this is usually a problem with numerical inversion methods. It was further concluded that the LAI retrieval could be done accurately with only one directional reflectance, while eight or more directional reflectances were required to accurately retrieve the LAD (leaf angle distribution) parameters. Kimes et al. (2002) inverted a complex 3D model (DART) for a wide range of simulated forest canopies. They compared a traditional inversion approach with the neural network method and found the ANN method to be more accurate. The ANN approach proofed to be computationally efficient when inverting a complex reflectance model and could be applied on a per-pixel basis. Udelhoven et al. (2000) inverted the INFORM model using hyperspectral DAIS data and found the estimated structural forest attributes to have reasonable values.

2.3.4 Chemical attributes

Assuming no constraints related to climatic conditions, water availability, nutrient availability or other potentially limiting factors, many of the biogeochemical processes in terrestrial ecosystems are related to the content of chemicals in leaves (Waring & Running, 1998). For instance, net primary production depends strongly on the availability of nitrogen (Vitousek, 1982). Litter decomposition is governed by the leaf C/N ratio (Melillo et al., 1982) and also by the leaf lignin content. Photosynthesis is largely controlled by chlorophyll pigments. Therefore, quantifying foliar chemicals within vegetation should help assessing the physiological status of vegetation, monitoring phenological characteristics, and estimating productivity (Blackburn, 1998).
Leaves are the most important plant surface interacting with solar energy. Besides canopy structure, the concentrations of the leaf chemical constituents (pigments, water, protein, cellulose, and lignin) shape the canopy reflectance spectra (section 2.2). Reflectance spectra can be recorded by airborne or spaceborne spectrometers and converted to information on foliar chemical constituents. This information can then be used to model forest ecosystem processes. Estimation of foliar biochemicals can be done at three different levels based on reflectance spectra either from dried leaves, fresh leaves, or canopies. An overview of literature results on the estimation of foliar chemicals from spectral variables using different methodologies.

Table 2.3: Selected literature results on the estimation of foliar chemicals from spectral variables using different methodologies. Chemical attributes: L = Lignin, N = Nitrogen, P = Protein, C = Cellulose, W = Water, DM = Dry matter, Chl = Chlorophyll; Level: D = Dried, ground leaves, F = Fresh, whole leaves, C = Canopy, G = Ground, A = Airborne, S = Synthetic; Methodology: MSR = Multiple stepwise regression, BLR = Bivariate linear regression, IRM = Inversion of reflectance model, ANN = Artificial neural network; Spectral variable: R = Reflectance, R' = First derivative of reflectance, diff = Reflectance differences at neighbouring bands, REIP = Red edge inflection point, BDN = Band depth normalization, VI = Vegetation index

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Species</th>
<th>Level</th>
<th>Methodology</th>
<th>Spectral Variable</th>
<th>Wavelengths [nm] and/or relation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>N, L</td>
<td>Deciduous</td>
<td>D</td>
<td>MSR</td>
<td>log(1/R)</td>
<td>various (N); various (L)</td>
<td>Peterson et al., 1988</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
<td>1240 (N); 1225, 2305, 2365 (L)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C (A)</td>
<td></td>
<td></td>
<td>none (N), 1500-1700 (L)</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>Various</td>
<td>C (A)</td>
<td>MSR</td>
<td>log(1/R)'</td>
<td>1256, 1311, 1555 (L)</td>
<td>Wessman et al., 1988</td>
</tr>
<tr>
<td>N ; L</td>
<td>Deciduous</td>
<td>C (A)</td>
<td>MSR</td>
<td>diff'</td>
<td>1265, 1555 (N); 1265, 1592 (L)</td>
<td>Wessman et al., 1989</td>
</tr>
<tr>
<td>Chl</td>
<td>Amaranthus</td>
<td>F</td>
<td>MSR</td>
<td>R'</td>
<td>710, 718, 724</td>
<td>Curran et al., 1992a</td>
</tr>
<tr>
<td>Chl</td>
<td>Sugar beet</td>
<td>C (G)</td>
<td>IRM</td>
<td>R</td>
<td>Good accuracy</td>
<td>Jacquemoud &amp; Baret, 1993</td>
</tr>
<tr>
<td>Chl, W</td>
<td>Sunflower</td>
<td>F</td>
<td>BLR</td>
<td>R-ratios</td>
<td>680, 430 (Chl); 979, 900 (W)</td>
<td>Penuelas et al., 1994</td>
</tr>
<tr>
<td>N, Chl,</td>
<td>Conifers</td>
<td>C (A)</td>
<td>MSR</td>
<td>R'</td>
<td>552, 1184, 2164 (N); 1069, 1683, 858 (Chl)</td>
<td>Johnson et al., 1994</td>
</tr>
<tr>
<td>L</td>
<td>Pine</td>
<td>D</td>
<td>MSR</td>
<td>log(1/R)</td>
<td>2180, 2230 (N); Validation on AVIRIS-spectra: r=0.74 (N), r=0.55 (L)</td>
<td>Gastellu-Etchegorry et al., 1995</td>
</tr>
<tr>
<td>P, C, L</td>
<td>Various</td>
<td>D, F</td>
<td>MSR</td>
<td>R', log(1/R)</td>
<td>5 λ; no typical wavelengths associated with absorption features; dry samples performed better than fresh material</td>
<td>Jacquemoud et al., 1995</td>
</tr>
<tr>
<td>Chl, C+L</td>
<td>Various</td>
<td>D</td>
<td>IRM</td>
<td>R</td>
<td>R²=0.84 (C+L); R²=0.50 (C+L); 0.65 (Chl)</td>
<td>Jacquemoud et al., 1996</td>
</tr>
<tr>
<td>DM</td>
<td>Various</td>
<td>D</td>
<td>IRM</td>
<td>R</td>
<td>DM (g per unit leaf area) is the only accurately retrieved parameter</td>
<td>Fourty et al., 1996</td>
</tr>
<tr>
<td>N, L, C</td>
<td>Various</td>
<td>D, F</td>
<td>MSR</td>
<td>log(1/R)', log(1/R)''</td>
<td>6 bands selected were not related to known absorption features; R² were not significantly higher than for randomized N-log(1/R)' datasets</td>
<td>Grossman et al., 1996</td>
</tr>
</tbody>
</table>
methodologies is given in Table 2.3.

## Dry leaf level

Remote sensing of chemical attributes started with spectrometric laboratory analysis of dried ground foliage. The methodology has been used to successfully derive protein, lignin, and other leaf constituents (Norris et al., 1976). Strong relationships between the concentration of foliar chemicals and reflectance were also found by other researchers for dry and ground leaves using multiple regression techniques (Card et al., 1988; Elvidge, 1990). More reliably results were obtained when normalised band depths were used instead of reflectance values (Kokaly & Clark, 1999; Curran et al., 2001). The inversion of a leaf reflectance model (PROSPECT) showed results of varying success. While Jacquemoud et al. (1996) obtained reasonable estimates on dry (but not on fresh) leaves, Fourty et al. (1996) were not able to
retrieve any of the single biochemical compounds within dry leaves; total amount of dry matter per unit leaf area was the only variable to be accurately retrieved.

**Fresh leaf level**

To bridge the gap between established laboratory methods and remote sensing, controlled experiments on fresh whole leaves have been undertaken by various researchers. However, the influence of water is strong throughout the near infrared region where most of the sought biochemicals have their absorption features. Consequently, leaf water can obscure absorption features related to foliar chemicals, such as protein, lignin, and cellulose in green leaf materials with considerable consequences for their estimation from reflectance spectra. For instance, Kokaly & Clark (1999) concluded that their method and those using log(1/R) derivatives would be sensitive to variations in leaf reflectance that may be solely caused by changes in leaf water content. Elvidge (1990) draws the conclusion, that remote sensing techniques for estimating canopy protein or canopy nitrogen content have to be based on indirect rather than direct measurements. Curran et al. (1992a) found good relationships between estimated and measured chemical concentrations for water, chlorophyll and starch (all $R^2>0.80$). A study undertaken by Fourty & Baret (1998) showed that the detailed biochemistry (protein, lignin, cellulose) was not accessible in an accurate way when observing a large range of species; they recommended working on a limited set of species. Grossman et al. (1996) found stepwise multiple regression relating chemistry data and derivative reflectance spectra to give relatively large coefficients of determination even for randomised data sets. The bands selected in their study lacked correspondence to those reported in other studies. Inversion of a reflectance model (PROSPECT) on the basis of fresh leaves worked out rather poor for cellulose and lignin, but reasonable for chlorophyll (Jacquemoud et al., 1996).

**Canopy level**

In the transition from the leaf to the canopy level, strong disturbing effects are introduced, such as canopy architecture (leaf area index, leaf angle distribution) and influence of soil background. Compared to controlled laboratory conditions, airborne or spaceborne studies bring in further disturbances, such as variable solar illumination and atmospheric conditions, along with instrument characteristics (signal-to-noise ratio, spectral bandwidth, geometric field of view). Together, these factors cause a number of problems in the usage of remote sensing data to the determination of foliar biochemistry. Despite these problems it has been attempted to derive chemical attributes related to forest canopies from the analysis of imaging spectrometry data (Peterson et al., 1988; Wessman et al., 1988, 1994a; Johnson et al., 1994; Gastellu-Etchegorry et al., 1995). Results suggested that spectroradiometric measurements of plant canopies can be correlated with variations in canopy chemistry, involving lignin,
CHAPTER 2 FUNDAMENTALS

nitrogen and cellulose concentrations. However, wavebands selected by multiple linear regression were often not consistent with the absorption features of the chemicals within the leaves. Furthermore, relationships from one forest site showed poor predictive performances when applied to another site (Grossman et al., 1996). Estimation of chlorophyll from remotely sensed imagery can generally be achieved more reliably than the other chemicals, as the prominent chlorophyll absorption feature is not masked by leaf water. For instance, Jago et al. (1999) related the red edge inflection point derived from hyperspectral image data to the mean chlorophyll concentration of grassland and winter wheat sites with moderate success. Various authors successfully estimated chlorophyll from leaf reflectance spectra (Curran et al., 1992a; Penuelas et al., 1994; Jacquemoud et al., 1996; Datt, 1998; Atzberger & Werner, 1998; Daughtry et al., 2000; Schlerf et al., 2003a) and canopy reflectance spectra (Jacquemoud & Baret, 1993; Blackburn, 1998; Demarez & Gastellu-Etchegorry, 2000) using different approaches. Since a large amount of leaf nitrogen is contained in chlorophyll molecules, attempts were made to estimate LNC from reflectances in the chlorophyll absorption feature or the red edge inflection point (Lilienthal et al., 2000; Lamb et al., 2002).

There are many other examples, where data acquired by imaging spectrometers have been related to foliar chemicals, with varying success (e.g., Kupiec, 1994; Kupiec & Curran, 1995; Gholz et al., 1996; Lucas et al., 2000; Zarco-Tejada et al., 2001; 2004). However, there is still a long way to go before canopy chemical concentration can be estimated reliably to drive ecosystem simulation models (Curran, 1994a; Wessman, 1994).

One important limiting factor to reliably estimate foliar chemicals is the signal-to-noise ratio of imaging spectrometry data. Of all the applications of imaging spectroscopy in recent years, the estimation of foliar chemical concentration has proved to be one of the most demanding in terms of signal-to-noise ratio (Peterson and Hubbard, 1992). Specifications given in the literature on required SNRs for chemical estimates vary considerably. Gholz et al. (1996) compared minimum SNRs from a laboratory study (Nitrogen 37:1, Cellulose 65:1) with the approximate SNRs at vegetation reflectance given by the Jet Propulsion Laboratory and assessed data provided by the AVIRIS sensor to be sufficient for the estimation of foliar chemical concentrations. The minimum required SNR for this purpose has also been estimated by Smith and Curran (1996). From modeling the SNR, they showed that the required SNRs at nitrogen absorption features were always the smallest due to close relationships of spectral reflectance to protein and nitrogen concentrations. They suggest that out of three chemicals studied (nitrogen, lignin and cellulose) nitrogen could be estimated most successfully with remotely sensed data. Inflight SNR for the AVIRIS sensor was also determined by Smith and Curran (1996). They concluded that at the beginning of the 1993 flight season the AVIRIS was close to the threshold SNR required to estimate accurately the concentration of foliar chemicals in vegetation canopies (from 1995 SNR was set to double) but conceded that the use of more than one waveband when converting reflectance to
chemical concentration might reduce the SNR requirement slightly. Nonetheless, the former statement questions the validity of results obtained from early studies using AVIRIS data for estimation of foliar biochemistry (e.g., Peterson et al., 1988; Wessman et al., 1988).
Chapter 3 Material

Remote sensing in the geosciences is dealing with many kinds of material such as study sites, sensors, field and laboratory data, and imagery. These aspects are described in this chapter including the radiometric and geometric correction of the HyMap data.

3.1 Study sites

The fieldwork for this research was conducted at two sites, Idarwald and Gerolstein, both located in south-western Germany (Figure 3.1). HyMap images were acquired at both test sites at points in time close to the field data acquisition. The study on the determination of forest cover type and structural attributes was carried out at Idarwald. The research on the estimation of foliar chemicals was conducted at Gerolstein test site. The reason for this split was, that Idarwald offers a great number of homogenous spruce stands with a wide range of age classes and structural attributes but a rather limited variability in foliar chemicals due to its spatially invariable underlying geology. The Gerolstein area, on the other hand, hosts spruce forest standing on soils from very poor to very rich in nutrient availability which gives reasons to expect a relatively large variability in foliar chemicals.

3.1.1 Idarwald

The Idarwald forest (49°45’N, 7°10’E) is located on the north-western slopes of the Hunsrück mountain ridge. It covers an area of about 7,500 ha. Climatically the area is dominated by oceanic influences with primarily westerly winds. Luv effects at the west side of the ridges cause a relatively high mean annual precipitation of about 800-1000 mm. Low stratus above the mountain ridges give rise to additional precipitation during the winter term. The mean annual air temperature is 6.5°C with a summer maximum in July and the minimum temperatures occurring in January. In the winter months frost and snowfall are frequent (Arbeitskreis Standortkartierung, 1985). Topographically the area is characterised by remnants of peneplains, and quartzite ridges reaching a height of up to 770 m above mean sea level. Geologically the area is underlain by Devonian quartzite (Taunusquarzit) and
Devonian shale (Hunsrückschiefer). Taunusquarzit consists to 98 percent of quartzitic sand grains cemented by silicic acid and is characterised by an enormous weathering resistance. Weathering products of Taunusquarzit are almost pure sands while the Hunsrückschiefer weathers to sandy loam. Loess sedimentation and solifluction during the Pleistocene had a major influence on the soil formation. The predominant soil types are rankers, podzols, and dystric cambisols (podsolige Braunerde). The soils range from well to poorly drained. They are generally low in nutrient status and are partially acidified (Reichert & Stets, 1980; Arbeitskreis Standortkartierung, 1985).

The dominant forest species are Norway spruce (*Picea abies* L. Karst.), beech (*Fagus sylvatica*), oak (*Quercus petraea*) and Douglas fir (*Pseudotsuga menziesii*). They cover 53, 18, 11, and 5 percent of the forested area, respectively. The predominant natural forest community in the area is *Luzulo-Fagetum typicum*, but afforestation starting at the end of the 18th century lead to a wide spreading of Norway spruce (Vohland, 1997; Arbeitskreis Standortkartierung, 1985).

Figure 3.1: Location of study sites. Data sources: ATKIS-Geobasisdaten (Landesvermessungsamt Rheinland-Pfalz), ESRI data (U.S. Geological Survey).
Besides natural factors, forest stand structure is largely determined by actual management practices. The timber forest or high forest (“Hochwald”) is a forest with trees that have evolved from seeds. There are two management types of high forest: forest with cutting areas, and single-tree selection forest (“Plenterwald”). Timber forest with cutting areas denotes a forested area that is subdivided into cutting areas (stands, divisions), in which the composition of the tree stock is rather uniform, especially in terms of age. Silvicultural measures comprise regeneration, tending, and thinning. Thinning is the extraction of some of the young trees in a forest so that the remainder grow and develop fully. The aim is to remove as much timber as possible while maintaining output to maximise the profitability of a site. Thinning has an major influence on the temporal LAI dynamic. At Idarwald, after a maximum LAI value at an age of 20-30 years is reached, stands of Norway spruce are usually thinned at an age of 30-50. After 60-70 years an increase in LAI may take place up to an age of 100 where logging can

Table 3.1: Soil substrates occurring at the Gerolstein test site (from Schwind & Schüler, 2001 and Schwind, personal communication). German denominations appear in brackets. For definition of the term “Soil substrate” refer to text.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Soil substrate</th>
<th>Underlying geology (epoch)</th>
<th>Typical soil type (FAO classification)</th>
<th>Base saturation</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLdu</td>
<td>Loam formed from quartzite, sandstone and shale (Verwitterungslehme unterdev. Quarzit-Sandstein-Tonschiefer Böden)</td>
<td>Lower Devonian (Unterdevon)</td>
<td>Eutric Cambisole (Skelettreiche Braunerde)</td>
<td>low-medium</td>
</tr>
<tr>
<td>wDLT</td>
<td>Wet loam cover above clay loam (Staunasse Decklehme über Tonlehm)</td>
<td>Lower Devonian (Unterdevon)</td>
<td>Planosole-Eutric Cambisole (Pseudogley-Braunerde)</td>
<td>medium</td>
</tr>
<tr>
<td>KVLdm</td>
<td>Loam formed from limestone and dolomite (Kalk- und Dolomitverwitterungslehme)</td>
<td>Mid- and Upper Devonian (Mittel- und Oberdevon)</td>
<td>Calcaric Cambisole, Rendzina (Kalkbraunerde, Rendzina)</td>
<td>very high</td>
</tr>
<tr>
<td>Ssm</td>
<td>Sands (Sande des Mittleren Buntsandsteins)</td>
<td>Mid-Buntsandstein (Mittlerer Buntsandstein)</td>
<td>Podzol (Podsol)</td>
<td>very low</td>
</tr>
<tr>
<td>DLsrm</td>
<td>Loam cover above sandstone (Decklehme über Sanden des Mittleren Buntsandsteins)</td>
<td>Mid-Buntsandstein (Mittlerer Buntsandstein)</td>
<td>Podzolic Cambisole (Podsol-Braunerde)</td>
<td>low</td>
</tr>
<tr>
<td>BVL</td>
<td>Loam formed from basalt (Basaltverwitterungslehme)</td>
<td>Tertiary-Quaternary (Tertiär-Quartär)</td>
<td>Eutric Cambisole (Braunerde)</td>
<td>high</td>
</tr>
<tr>
<td>DLB</td>
<td>Loam cover above basalt (Decklehme über Basalt)</td>
<td>Tertiary-Quaternary (Tertiär-Quartär)</td>
<td>Eutric Cambisole (Braunerde)</td>
<td>high</td>
</tr>
<tr>
<td>Kol+</td>
<td>Colluvium (Kolluvium)</td>
<td>Quaternary (Quartär)</td>
<td>Gleysole-Cambisole (Gley-Braunerde)</td>
<td>high</td>
</tr>
</tbody>
</table>
lead to gaps in the canopy cover. Depending on the application of thinning measures, maximum variation of LAI can be expected at age 30-70 (personal communication Wagner, Womelsdorf; cited from Vohland, 1997).

Another method of regeneration is the clear cutting system, which involves clearing an entire area and then replanting it in a uniform and homogeneous fashion. Nowadays, clear cutting measures are restricted by law requiring special permission by the forest authority of the respective federal state if they exceed a certain size. By forest management in cutting areas, the vast majority of the forests in Germany (approximately 97%) have been converted into one-age-cohort plantations. This particular form of management is now considered outdated, but still largely shape the present-day appearance of our forests due to long time-lags (Häusler & Scherer-Lorenzen, 2001).

Soils formed over the quartzite ridges typically show a low buffer capacity as a consequence of the low nutrient availability. As a consequence, soil acidification has become a major problem in the study area during the 80s and early 90s of the last century. At the beginning of the 1990s, measurements in the upper soil layer revealed pH values as low as 3.8, indicating that protons are buffered under release of aluminium ions out of clay minerals (Block et al., 1991). Free Al$^{3+}$-ions cause damage to plants’ fine roots and the mykorrizha. High soil acidity may initiate additional processes: i) release of heavy metal ions, further increasing aluminium toxicity; ii) wash out of cation bases, such as Ca$^{2+}$, Mg$^{2+}$, K$^+$ and Na$^+$ resulting in a low cation exchange capacity. From the latter process, a limitation of the plant nutrient supply can be expected (Block, 1993).

Nutrient concentrations from chemical analyses of foliage samples can give evidence towards restricted nutrient supply. Insufficient supply with magnesium (Mg), indicated by yellowing of leaves or needles as a consequence of chlorophyll decomposition has been observed particularly at the mountain ridges of the Idarwald during the 1980s. From 1990 onwards, intensive fertilization measures (3-6 tons of dolomite per hectare) have been carried out at several forest stands in order to increase their Mg-supply and neutralize acids present in the

<table>
<thead>
<tr>
<th>Soil substrate</th>
<th>Soil profile</th>
<th>CEC$_{eff}$ [mmol kg$^{-1}$]</th>
<th>BS [%]</th>
<th>N [mg kg$^{-1}$]</th>
<th>C/N of Ah</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vldu</td>
<td>Ah-Bv-Cv</td>
<td>138-52-39</td>
<td>27-5-4</td>
<td>4310-700-380</td>
<td>24</td>
</tr>
<tr>
<td>KVLDm</td>
<td>Ah-Bv-Cv-IICv</td>
<td>109-307-482</td>
<td>75-97-98</td>
<td>3480-1300-650</td>
<td>12</td>
</tr>
<tr>
<td>Ssm</td>
<td>Aeh-Bv1-Bv2-Cv</td>
<td>105-83-26-6</td>
<td>8-3-8-27</td>
<td>2320-1230-200-100</td>
<td>26</td>
</tr>
</tbody>
</table>
soil. As a result, the portion of yellowed trees decreased from 40 percent in 1980 to just 2 percent in 1995 (Ministerium für Umwelt und Forsten Rheinland-Pfalz, 1995). Therefore, yellowing is not anymore considered as a major problem at Idarwald.

Due to the limited spatial variability in soil nutrient availability as a consequence of the relatively uniform substrate characteristics at the Idarwald test site, spatial variability is also low in leaf chemical concentrations. However, remote sensing of foliar biochemistry requires a location where forest grows on both rich and poor sites. Therefore, another test site was established which is characterised by a great small-scale variability in lithology.

3.1.2 Gerolstein

The Gerolstein test site (50°15’N, 6°40’E) is located in the central part of the Eifel mountain. The mean annual precipitation is about 800-900 mm; the mean annual air temperature is 7°C.

The manifold soil types are a product of the extreme diversity of bedrock types that occur in the area. Geologically the area is underlain by lower Devonian quartzite-sandstone-schist formations; middle Devonian limestones (calcite and dolomite), Triassic sandstones, as well as Tertiary and Quaternary basalts. In parts of the area the underlying bedrock was mixed with Pleistocene loess cover or periglacial layers. Small river valleys are covered by colluvial material that has been washed off the slopes and deposited at the foot slopes during the Holocene (Schwind & Schüler, 2001). To classify the complexity of the soil types, the concept of “soil substrates” is used in Germany by the Federal Research Centre for Forestry and Forest Products. The term “soil substrate” (german: Bodensubstrat) describes in a combined manner the grain-size distribution of the soil material, the type of bedrock and the stratification. Eight different soil substrates were identified at the study site (Table 3.1). Despite the extreme range in nutrient availability (Table 3.2), stands of Norway spruce forest are found on all eight substrate types.

3.2 Field and laboratory data

3.2.1 Field data

One way to determine certain characteristics of an entire forest is to measure every single tree in it. For practical reasons, measurements made of some individual trees in the forest are scaled up to estimate what is available from the whole forest. This process of ‘scaling up’ involves two steps. Firstly, measurements of individual trees are made in stands. A stand refers to a more or less homogenous group of trees in a forest. Stand measurements are usually performed on a per unit ground area basis, for instance the number of stems per hectare. Stand measurements can also be recorded as an average for the stand, for example,
the average of tree stem diameters. The second step in scaling up means that many stands scattered throughout the forest are measured to estimate the total amount or the distribution of an attribute across the entire forest (West, 2004).

Figure 3.2: General map of Idarwald test site with stand boundaries and locations of test plots.
Idarwald test site

In 2000, 42 relatively homogenous Norway spruce stands were identified at the Idarwald study area (Figure 3.2). Dolomite had been added to 25 of the stands between 1986 and 1995. Within each stand, a square 0.09 ha plot was established. The central location of each ground plot was determined with an accuracy of about ±5 m using a differential global positioning system (GPS). Signal distortion within the forest stands did not give as high accuracies as specified by the manufacturer. From a total of 42 stands, one was not covered by the HyMap imagery, and another one was thinned in the period between ground data collection and the HyMap overpass. These two stands were excluded from the analysis reducing the dataset to 40 stands.

At each plot, measurements of forest structural attributes were carried out. Attributes that have been measured include leaf area index (LAI), stem density (SD), canopy cover (CO), perimeter-at-breast-height (PBH), stand height (H) and crown diameter (CD). LAI was estimated using a Li-Cor LAI-2000 Plant Canopy Analyser. The LAI-2000 estimates effective LAI using measurements of diffuse solar radiation above and below the forest canopy. Several factors such as sky conditions, topography, foliage clumping, woody materials, and plant phenology all affect LAI estimates (Fournier et al., 2003). The LAI-2000 was only operated under overcast sky conditions during 10-16 hours daytime. A 270-degree view restrictor was used on the sensor. Within each plot, below canopy measurements were taken at 10 regularly spaced points from which the average was calculated. As no second device was available to operate the LAI-2000 in a two-sensor mode, above the canopy measurements were taken in a nearby open field before entering the plots. It was carefully paid attention to stable sky conditions between the open field and plot measurements. None of the outer rings were eliminated in the gap-fraction inversion. Despite the non-random distribution of leaves, no corrections for shoot level clumping and stand level clumping were applied. Also no contributions of woody surfaces were subtracted as the influence of woody components on the measurements is difficult to assess. It was assumed that the underestimation of LAI due to clumping effects was somehow compensated by the overestimation of LAI through woody structures (Fournier et al., 2003). As no corrections were applied, the retrieved LAI-2000 measurements represent an effective plant area index instead of the real leaf area index (see section 3.2.1). In the following sections these measurements are named effective leaf area index and abbreviated as LAI. Stem density (SD) or stocking density, the number of tree stems per unit area, was obtained by counting the number of trees in a plot and dividing the number by the plot area measured in hectare. Canopy cover (CO) or fraction of canopy closure, was
visually estimated in steps of 5 percent. Perimeter-at-breast-height (PBH) is defined as the trunk perimeter at 1.3 m above the forest floor. It is equivalent to diameter-at-breast height.

Figure 3.3: General map of Gerolstein test site with stand boundaries, locations of field plots and locations of targets used for radiometric calibration.
(DBH) which is often used instead. PBH was measured for 10 randomly selected trees within each plot and the stand average was calculated. Stand height (H) was calculated from the mean height of three dominant trees that were randomly selected within each plot. The height of each tree (from the ground to the top) was determined from angular measurements. As forest stands at Idarwald consist of trees of the same age, dominant trees within a stand have similar heights. Hence, the selection of three trees to represent H was considered to be appropriate. Crown diameter (CD) was determined as the mean of maximum and minimum crown diameters. From five representative trees per plot the average was calculated. CD was only measured in forest stands older than sixty years (n = 28).

Assuming that the volume of a single stem can be approximated by a cone, a new variable stem biomass (SBM, unit: t ha⁻¹) was calculated from PBH, H, and SD:

\[
SBM = \frac{\pi \cdot \left( \frac{PBH}{2 \cdot \pi} \right)^2 \cdot H}{3} \cdot SD \cdot \delta_{WOOD}
\]

where \( \delta_{WOOD} \) is the density of fresh wood (for Norway spruce: 0.47g/cm³ at 12-15% moisture content; Sedlmayer, 2004).

Figure 3.4: Woodman climbing a tree at Idarwald test site (March, 2000).
Gerolstein test site

In August 2003, 13 relatively homogenous Norway spruce stands were identified at the Gerolstein test site (Figure 3.3). Three trees were randomly selected within each stand. Skilled woodsmen climbed the trunks up to the treetop and cut off three branches from the outer part of each crown (Figure 3.4). Foliage from the sampled branches was removed separately by age class (1st and 3rd year) according to their position on the branch, bagged, labelled and placed in cold boxes in the field. As freezing and thawing are known to change the spectral properties of leaves and needles, the foliage samples were stored in refrigerators at +5°C after the transport to the geosciences laboratory of the University of Trier. In total, 78 samples were obtained (13 stands x 3 trees x 2 age classes). From these samples sub-samples were randomly selected and subjected to standard wet chemical analysis to obtain measures of the concentration of chlorophyll a and b, water, nitrogen, and other chemical elements (section 3.2.2). Additional sub-samples were randomly selected for reflectance measurements (section 3.2.3). To validate predictive models developed on the data obtained at Gerolstein (section 7.4.4) foliar samples from Idarwald were used. At Idarwald test site samples had been collected during March 2000 in a same fashion as previously described for the Gerolstein area. In total, 30 samples were obtained (5 stands x 3 trees x 2 age classes). Chemical and spectrometric measurements and pre-processing steps were identical for both sets of data.

3.2.2 Chemical laboratory analysis

To determine leaf water concentration (C_W), approximately 5 g of fresh sample were weighted, dried in an oven at 105 °C to constant weight (24 hrs), and were then re-weighted. Leaf water concentration was calculated as (fresh weight [g] – dry weight [g]) / dry weight [g] x 100.

Chlorophyll concentration (C_AB) of the foliage was determined as described by Lichtenthaler (1987). Approximately 0.5 g of fresh sample were weighted and ground by hand in acetone with aid of quartz sand until no green colour was left in the residual material. After the chlorophyll solution was decanted off the residue, filled up to a volume of 50 ml with acetone, centrifuged for 30 minutes, it was measured in a laboratory spectrophotometer. Chlorophyll a was measured at its absorption maximum of 663 nm and chlorophyll b at 646 nm. Concentration of chlorophyll a+b (C_AB) was converted from unit “grams per litre” to “milligrams per gram of dry weight”: C_AB [mg g⁻¹] = C_AB [g l⁻¹] x volume [l] x 1000 / dry sample weight [g]. The extraction of chlorophyll was carried out under cool and totally dark conditions as leaf pigments tend to decompose when exposed to light.
Total nitrogen concentration \((C_N)\) of the foliage was determined through CHN elementary analysis (Wilson, 1990). The samples are broken down into its atomic components and are then separated. To break the sample down it is combusted in an oxygen atmosphere at 980°C. At this temperature all of the elements to be detected react with oxygen to form \(\text{CO}_2\), \(\text{H}_2\text{O}\), and \(\text{N}_2\text{O}_y\). The gas is then led over copper granulate material where the nitrogen oxides are reduced to \(\text{N}_2\). Next, the gases are carried via a He-stream to a detector. The detector reports a value to the computer which compares it to the known value of a standard. These values are calculated based on the sample weight. Results are reported in percent element by weight.

### 3.2.3 Spectral reflectance measurements

The directional-hemispherical spectral reflectance and transmittance of plant leaves are typically measured by laboratory spectroradiometers equipped with an integrating sphere to average the signal reflected or transmitted in all directions. Problems typically occur with needles as their size is generally shorter than the diameter of the sample port of integrating spheres. As a consequence, alternative techniques consist of measuring the reflectance of a mat of needles arranged side-by-side into a single layer or measuring the reflectance of an infinite number of needles, e.g. a layer stack (Despan & Jacquemoud, 2004).

Reflectance spectra were measured using a Field-Spec-II spectroradiometer under laboratory conditions. The target was illuminated by a 1000 W halogen light source at a zenith angle of 45° and a distance of approximately 60 cm. The optical head of the Field-Spec-II was positioned at nadir (0° viewing angle) at a distance of 10 cm above the target. It was made sure that the instrument’s field of view lay well within the target’s perimeter. The target consisted of needles arranged in an optical dense layer of about 1 cm thickness over a matt black background. Relative spectral radiances between 350 and 2500 nm were recorded at

![Figure 3.5: Signal-to-noise ratio measured while viewing a 50% reflectance target at a solar zenith angle of 30 degrees; the pixel integration time was equivalent to operation at 14 lines per second (from Cocks et al., 1998).](image)
1 nm steps in 2151 spectral wavebands. Each sample measurement comprised 50 single measurements of the target and was repeated four times thereby rotating the target to average out any possible directional effects. The reference panel (Spectralon) was measured before and after the target measurements. The absolute spectral reflectance ($\rho_{\text{target}}$) for each waveband was calculated from the radiance counts of measurements of the target (DN$_{\text{target}}$) and the reference (DN$_{\text{reference}}$) and the reflectance of the reference ($\rho_{\text{reference}}$) as $\rho_{\text{target}} = \frac{\text{DN}_{\text{target}}}{\text{DN}_{\text{reference}}} \times \rho_{\text{reference}}$. The 5 spectra recorded for each sample were averaged to produce a single spectrum. The resulting spectra were smoothed using a 31 nm wide moving Savitzky-Golay filter to reduce instrument noise.

### 3.3 HyMap image data

The estimation of quantitative forest properties from remote sensing data requires acquisition and preprocessing of image data. This section outlines the design and operation of the Hyperspectral Mapper (HyMap) and describes the radiometric and geometric correction that has been applied to the HyMap data.

#### 3.3.1 The HyMap sensor

HyMap was designed and built by Integrated Spectronics, Australia, and operated by HyVista. It has become operational for scientific and commercial tasks in 1998. The spatial configuration is as follows: The instrument collects data in across-track direction by mechanical scanning and in along-track direction by movement of the airborne platform. Its instantaneous field of view (IFOV) is 2.5 mrad along-track and 2.0 mrad across-track and the field of view (FOV) is 60 degrees (512 pixels). Typically, the sensor is flown aboard a light, twin engined aircraft at altitudes of 2000-5000 m above ground level. The swath width at these altitudes is 2.3 to 4.6 km and the ground resolution 5 to 10 m (along-track) (Cocks et al, 1998).
The spectral configuration of HyMap provides contiguous sampling across the spectral range of 0.45–2.48 μm except for gaps in the strong atmospheric water absorption bands near 1.4 μm and 1.9 μm. Additionally, it collects broadband information in the thermal infrared domain. A complete spectrum over the range of 0.45–2.48 μm is recorded in a sampling interval of 13–17 nm by 4 spectrographic modules (Table 3.3). The visible module includes a silicon (Si) detector array operated at ambient temperatures and the near-IR, SWIR1, and SWIR2 modules incorporate indium-antimonide (InSb) arrays cooled by liquid nitrogen. Each spectrographic module provides 32 spectral channels giving a total of 128 spectral measurements for each image pixel (Cocks et al, 1998).

Airborne imaging spectrometers have to cope with a relatively weak illumination source (the sun) compared to laboratory illumination sources, sensor-to-object distances of several kilometres, and very short dwell times. Together, these factors suggest that imaging spectrometers generally have a low signal-to-noise ratio (SNR) (Curran & Dungan, cited from Kupiec, 1994). To meet the requirements for the detection of subtle variations in the spectral properties of image objects, sensor technology has been driven to achieve higher SNRs. Ground measurements by Cocks et al. (1998) have shown that in the spectral regions of the atmospheric windows, the HyMap sensor achieves SNRs approaching 1000:1 or better (Figure 3.5). These ground based measurements of SNR are a very good indicator of the sensor’s in-flight performance.

Figure 3.6: Raw (left) and cross-track illumination corrected (right) HyMap image profiles (upper part) and corresponding portions of the HyMap scene (lower part) at Idarwald. The yellow lines in the images indicate the course of the profiles. Profiles are shown for the red (Bd. 15), near-IR (Bd. 34), and mid-IR (Bd. 83) domains. Images are displayed at band combination 34-83-15 (RGB).
3.3.2 The HyMap data sets

HyMap was flown over Idarwald test site at 17\textsuperscript{th} July 1999. The data were recorded at 13:00 hrs CET at an average flying height of 1980 m above ground level. The resulting ground resolution was about 5 m with a full scene covering approximately 3 km x 10 km. The flight line was run in a NE–SW direction parallel to the mountain ridge. The acquired single scene was free of cloud cover. A number of 14 HyMap spectral channels with high noise were identified as bad bands and removed from the dataset. After removing these channels, a total of 114 wavebands remained out of the original 128. All subsequent data analyses were carried out on the 114-band dataset.

HyMap was flown over Gerolstein test site at 14\textsuperscript{th} July 2003. The data were recorded at 12:40 hrs CET at an average flying height of 3400 m above ground level. The resulting ground resolution was about 7 m with a full scene covering about 4 km x 13 km. The flight line was run in almost N-S direction. The acquired scene was free of cloud cover. A number of 16 HyMap spectral channels with high noise were identified as bad bands and removed from the dataset. After removing these channels, a total of 110 wavebands remained out of the original 126. All subsequent data analyses were carried out on the 110-band dataset.

3.3.3 Radiometric correction

Radiometric correction of image data is exceptionally important to remote sensing of coniferous forests, as the small reflectance signal generated by conifer canopies is strongly influenced by terrain and atmospheric effects (Peterson & Running, 1989). When quantitative attributes are to be estimated from remote sensing data, surface reflectance is the preferred or

---

Figure 3.7: Schematic representation of the sensor recalibration.
required data type (Peddle et al., 2003). Radiometric correction of the HyMap data set acquired over both test sites consisted of i) an across-track illumination correction and ii) additional radiometric correction involving atmospheric correction, correction for illumination effects caused by topography, and in-flight calibration of the HyMap sensor system. The vicarious calibration of the instrument implies the computation of updated calibration coefficients (illustrated in Figure 3.7). From reflectance measurements collected at several reference targets on the ground, at-sensor radiance is modelled using a radiative transfer code. Modelled radiance is then related to measured image radiance through a linear regression to obtain a set of updated in-flight calibration coefficients.

Atmospheric correction of the HyMap data sets was achieved with an in-house developed software package (AtCPro 3.01) which is originally based on the formulation of radiative transfer in 5S and 6S (e.g., Tanré et al., 1990) (section 2.1.2). This code has been modified to account for atmospheric extinction processes as a function of sensor and terrain altitude (Hill & Sturm, 1991). It has been further extended for the ability to estimate the aerosol optical thickness directly from dark objects in the satellite image (Hill, 1993). A topographic correction module has been integrated where targets are assumed to have Lambertian reflectance characteristics (Hill et al., 1995). Recently, the method has been extended to cope with specific requirements of airborne hyperspectral image data by adding a module that produces spatially distributed maps of atmospheric water vapour to be included into the atmospheric correction (Hill & Mehl, 2003).
HyMap at Idarwald test site

Across-track illumination correction was applied to correct for the effects of change in sensor view angle that typically affect sensors with a large total field of view (Kennedy et al., 1997). To determine the magnitude of these effects, values of digital numbers (DN) representing conifer forest were plotted against image pixel number in the across scan direction. A systematic increase in values of DNs with view angle in the down-Sun viewing direction and a decrease in the opposite-Sun viewing direction was observed (Figure 3.6). This well known trend can be attributed to the effect of anisotropy and canopy shadow. As most vegetation canopies show a backward scattering characteristic the opposite-Sun side of the scanned image receives less reflected energy than the down-Sun side (Hildebrandt, 1996; section 2.2.4). To correct this view angle effect an across track illumination correction was applied to each spectral band independently (ENVI 3.4, Research Systems). For this purpose, a second-order polynomial was fitted to the data. Based on the fitted polynomials, a normalization procedure was applied.

The atmospheric correction of the HyMap data set used additional field data collected for a specific calibration flight performed during the same day in a different area only 30 km away from the Idarwald study site. There, measurements of atmospheric beam transmittance were conducted with a CIMEL photometer during the time of both flights. Their evaluation yielded an aerosol optical depth of $\tau_a = 0.33$ at $\lambda = 0.55 \mu$m representing a horizontal visibility of approximately 15 km. Contemporarily, bi-directional reflectances of several reference targets with different surface characteristics (gravel, asphalt, grass and water) were measured at the ground with an ASD Field-Spec-II instrument. From these ground measurements, together with the derived aerosol optical depth and a water vapour concentration of 3.8 g cm$^{-2}$ that was obtained from an iterative approximation based on several MODTRAN runs, the at-sensor radiance was modelled. Modelled radiance and measured digital numbers from the image were then related through a linear regression to obtain a set of updated HyMap in-flight calibration coefficients. It turned out that, except for single noisy bands with limited radiometric performance, the resulting adjustments remained within ± 15 % of the HyMap pre-flight calibration values. Using these updated calibration coefficients together with estimates of the water vapour concentration it was possible to reconstruct a spatially differentiated water vapour map for the HyMap scene of Idarwald. Water vapour concentration for each image pixel was estimated from the depth of the water absorption features at 0.94 µm and 1.14 µm. In a final run, the water vapour map was integrated into the atmospheric correction processing of this image. This correction also included a correction for terrain induced illumination effects, for which specific DEM-derivates (slope, aspect, and the proportion of the visible hemispherical sky at each pixel) had been transferred to the geometry of the original HyMap image using the parametric image processing software PARGE developed by Schläpfer et al.(1998). The obtained reflectance spectra (Figure 3.8) point...
towards a good data quality and indicate that a sound foundation has been laid for subsequent quantitative data exploration using empirical or physical based methods (Hill, 2002).

**HyMap at Gerolstein test site**

The radiometric correction procedure of HyMap imagery acquired over Gerolstein test site was similar to the procedure applied to the Idarwald image that was described in the previous section. Just a few aspects are mentioned where the procedure differs from the one applied to the Idarwald image.

To correct for across track illumination effects, a first order polynomial was fitted to the data followed by a normalisation procedure. The aerosol optical depth was derived from estimates of the horizontal visibility that was provided by three meteorological stations located in the region (mean horizontal visibility: 35 km). At the day of the overflight, several reference targets were measured at the ground with an ASD Field-Spec-II spectroradiometer. The targets covered a wide range of surface characteristics and included short grass (football ground), long grass (pasture), cereals (grain field), asphalt (parking lot), calcareous material (mining dump) and water (pond). From these ground measurements, together with the derived aerosol optical depth and the iteratively derived water vapour concentration the at-sensor radiance was modelled. Modelled radiance and measured digital numbers from the image were then related through a linear regression (Figure 3.9) to obtain a set of updated HyMap in-flight calibration coefficients. The relations between measured and modelled radiances lie close to the 1:1 lines and have values of $r^2$ between 0.91 and 0.98 with best relations in visible and near-IR wavebands. The resulting adjustments remained within ± 10 % of the HyMap pre-flight calibration values except for single noisy bands. The obtained reflectance spectra exhibited an excellent data quality which is an important pre-requisite for mapping of foliar chemicals.
Figure 3.9: Modelled against measured radiance in six wavebands based on ten reference targets at Gerolstein test site.
3.3.4 Geometric correction

The image data were geometrically corrected to the local coordinate system (Gauss Krüger, Zone 2, Ellipsoid: Bessel 1841, Datum: Potsdam) using the parametric geocoding software PARGE (Schläpfer et al., 1998). For this purpose, the required DEM with an original pixel size of 20 m was resampled to the image pixel size (Idarwald: 5 m; Gerolstein: 7 m) using bilinear interpolation. In the case of the Idarwald image, 30 ground control points (GCPs) measured in the field by differential GPS were used to calibrate the in-flight auxiliary data. For the Gerolstein image, 53 GCPs were obtained from an orthophoto mosaic. Image resampling was performed using triangulation coding from centre pixels (Schläpfer et al., 1998). The root mean squared error of the geometric correction was 4.3 m in x-direction and 4.8 m in y-direction (Idarwald image) and 4.9 m in x-direction and 5.6 m in y-direction (Gerolstein image). With an image pixel size of 5 m and 7 m, respectively, the geometric correction of both data sets was accurate to within a pixel.

Figure 3.10: Extraction of spectra for various radii. Circle segments were obtained through buffering around the GPS-measured central plot position and subsequent masking using the conifer forest mask. Image pixels representing shaded spruce crown appear in a very dark green tone, sunlit spruce crowns appear in dark green, beech in bright green and road in purple tones.
3.3.5 Extraction of reflectance spectra from HyMap imagery

Idarwald image

Reflectance spectra have been extracted from the image data for all 40 forest stands under investigation. To collect representative image spectra in each plot, average reflectance spectra were computed from image pixels within certain radii around the central plot position. Average spectra were derived from radii of 7 m, 12 m, 17 m, 22 m, 32 m, 42 m, and 52 m. Pixels not representing spruce forest were cut out from the circles through a GIS operation resulting in circle segments (Figure 3.10). Mean and standard deviation of spectral reflectance were extracted. Mean reflectance was plotted against wavelengths for a typical forest stand for the near-IR wavelength region (Figure 3.11A) and the visible wavelength region (Figure 3.11B). It is evident that a critical value of reflectance is reached with increasing radius and stabilisation occurs. The plateau where reflectance remains constant lasts from a radius of 32 m until 52 m in both the visible and the near-infrared wavelength region. Therefore, the central position of the plateau at a radius of 42 m was considered to be the critical size from where local heterogeneity is not further increasing. The other forest stands essentially show similar patterns. Sometimes, values of reflectance plateau already at a smaller radius, but

Figure 3.11: (A) Change of reflectance with increasing circle radius across the visible region (HyMap-bands 4-12) and (B) the near-infrared region (HyMap-bands 20-32). (C) Change of the coefficient of variation with increasing circle radius across the visible region (HyMap-bands 4-12) and (D) the near-infrared region (HyMap-bands 20-32).
never a decrease in reflectance with increasing circle size is observed. A sharp decrease of reflectance in the near-IR region would indicate that the boundary of the stand is crossed, incorporating gaps in the forest cover. A sharp increase of reflectance in the near-IR region would indicate incorporation of broad-leaf forest which shows a considerably higher reflectance at the near-IR plateau (Figure 3.10).

In a similar fashion, the coefficient of variation (standard deviation of spectral reflectance divided by the mean spectral reflectance) was plotted against wavelength for the visible (Figure 3.11C) and the near-infrared (Figure 3.11D) region. As can be seen, the lowest variation occurs at radii of 42 m for both wavelengths regions.

**Gerolstein image**

As many of the investigated forest stands at Gerolstein test site were relatively small in relation to the image pixel size, a systematic investigation of the number of extracted pixels was not feasible. Instead, reflectance spectra were extracted from the image using the region growing tool under the image processing software ENVI (version 4.0). At each field plot in the image, four pixels were defined as a “region of interest” (ROI). Growing of ROIs was done in 4 directions (neighbours) with a standard deviation multiplier of 1 resulting in ROIs of 8 to 31 image pixels. From the selected image pixels mean reflectance spectra were computed and used for further analysis.

**3.4 Forest Geographic Information System**

A Forest Geographic Information System (FoGIS) was available for Idarwald test site. It contained the current forest inventory information (cutoff date: 01.10.1994) as had been made available by the regional forest authority (Forstdirektion Koblenz). The forest inventory data listed for each forest type present in a particular forest stand information about tree species, age, form of mixing, productivity, etc. This information was aggregated into a stand average in order to link attribute data to geographic data (stand polygons) by a 1:1 relation (Vohland, 1997). With the help of FoGIS it was possible to find suitable training and validation areas for image classification (Chapter 5) and to identify useful plots for field data sampling (section 3.2.1).
Chapter 4  Statistical and physical methods

4.1 Overview

Generally, it can be distinguished between empirical-statistical and analytical techniques for the derivation of leaf and canopy properties from reflectance data. The following main approaches exist for the utilization of remotely sensed spectral information to identify and characterize plant canopies (Ustin et al., 1999; Baret & Jacquemoud, 1994):

- Data analysis to determine the spatial extent of cover types based on their spectral response. Mapping of cover types is typically achieved through image classification. In classification, a statistically based decision rule is applied to categorize all pixels in an image to one of several cover classes (Lillesand & Kiefer, 2000). Image classification techniques (section 4.2) were used to map tree species and age classes.

- Empirical models that relate the spectral measurement through regression analysis to the variable of interest. Regression techniques comprise bivariate regression, multiple stepwise regression, partial least squares regression, and others. The spectral measurement is often manipulated in such a way, that it can be related to a given canopy property (Verstraete & Pinty, 1996). These so-called vegetation indices (VIs) include band ratios, spectral derivatives, spectral shifts, and others. One objective of using VIs is to enhance spectral features while minimizing confounding factors, such as soil optical properties or atmospheric effects (Demetriades-Shah et al., 1990). Empirical techniques are described in section 4.3.

- Analytical models. Radiative transfer models describe the interaction of radiation with leaves and vegetation canopies based on physical principles. Such a leaf or canopy reflectance model relates the spectral information via a complex mathematical function to leaf and canopy variables (Goel, 1988). A good model will simulate a realistic spectral signature from a given set of variables. Once a model is developed, estimates of these variables can be retrieved from reflectance spectra through inversion (Goel, 1989). One of the main problems with model inversion is the so-called ill-posed problem. It states that a given reflectance signal could be generated by more than one set of combinations of variable values, because different variables may affect the canopy reflectance in the same way. Consequently, a unique set of parameter
combination would not exist (Combal et al., 2000). Physically-based techniques are described in section 4.4.

- Spectral mixture analysis can be considered a simple reflectance model. In spectral mixture analysis, the reflectance spectrum of a given target is considered as composed of a small number of elementary objects called end-members. The (linear) combination of single end-member spectra is determined that describes best the observed spectrum. It provides an estimate of the relative coverage of the elementary surface types (e.g., vegetation) in the observed pixels. However, this approach assumes that radiative transfer processes involved are additive which is not valid for dense vegetation canopies (Smith et al., 1994). Therefore, spectral mixture analysis was not used in this research work.

The structural and chemical properties of a vegetation canopy control its spectral signature (section 2.2) and can in general be derived from remote sensing (section 2.3). The reflectance \( \rho_\lambda \) of a vegetation canopy at wavelength \( \lambda \) is determined by the vegetation parameters \( V \) plus more or less variable side conditions \( S \):

\[
\rho_\lambda = f(V_1, \ldots, V_n, S)
\]

To find the function \( f \) that relates the parameters of interest to the recorded signal is the fundamental goal in quantitative remote sensing (Verstraete et al., 1996). However, the side conditions comprise a number of variables that describe the atmosphere, illumination and viewing geometry, environmental conditions, etc. that are very difficult to measure. Therefore, finding \( f \) is not trivial. It is the overall objective of this research to compare different methods that can find transfer functions that relate remote sensing data of certain spectral properties to various forest attributes in order to enable their reliable estimation from operational sensors in the future.

The following sections describe general aspects of the methodology that were used in this research work. Information on how the methods were implemented in detail are given in Chapter 5, Chapter 6 and Chapter 7 as processing strategies and results are often strongly interrelated.

### 4.2 Image classification

#### 4.2.1 Introduction

Classification has been an important tool in digital image analysis for land resources applications since early Landsat missions when multispectral image data was regarded as a composition of multivariate measurement vectors for each image pixel. In classification, the
thousands or millions of such vectors that form an image are treated as class descriptors, and the spectral bands as explanatory variables related to these classes. The underlying assumption is that each class exhibits a unique multi- or hyperspectral signature. In the classification process, image pixels with similar signatures are grouped into classes which are pre-defined in supervised classification approaches. The result is a simplified image where pixels are labelled with their appropriate class type (Franklin et al., 2003).

Therefore, classification is the most logical approach for the prediction of the class membership of an image pixel when categorical attributes (e.g. forest type, tree species) are mapped. Yet, classification can also be used to map continuous forest attributes, such as cover, biomass or stand age, when those attributes are divided into ordinal categories (Franklin et al., 2003). For instance, the attribute ‘stand age’ with unit years can be divided into distinct age classes.

### 4.2.2 Spectral Angle Mapper

A number of classification methods exist that have been used in forest mapping from remote sensing data (Franklin et al., 2003). Most often described in the literature and implemented into imaging processing software packages are Minimum Distance, Box or Parallelepiped and Maximum Likelihood Classifier. These classifiers were all designed to cope with data of multispectral type. Other, more advanced classifiers, include Artificial Neural Networks and Decision Tree Classifiers. The availability of hyperspectral data has led to the development of new classifiers that could better handle large numbers of wavebands. For instance, the Spectral Angle Mapper (SAM) determines the spectral similarity between two spectra by calculating the angle between the spectra, treating them as vectors in a space with dimensionality equal to the number of bands. SAM compares the angle between the spectrum vector of the known class and each pixel vector (unknown class) in n-dimensional space. In the classification stage, the class with the smallest angle is assigned to the corresponding image pixel. Because this method uses only the direction and not the length of the vector, it is insensitive to illumination (Kruse et al, 1993). The angle $\alpha$ between the test spectrum $t$ and the reference spectrum $r$ is calculated as

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{n} t_i r_i}{\left(\sum_{i=1}^{n} t_i^2\right)^{0.5}\left(\sum_{i=1}^{n} r_i^2\right)^{0.5}}\right)$$  \hspace{1cm} (4.2)$$

where $n$ is the number of spectral bands.
4.3 Statistical methods

Empirical quantitative analysis in spectroscopy and in other scientific fields is based on the assumption that the amounts or concentrations of the constituents of interest in the samples are in some way related to the spectral measurement of those samples. The aim is to find a calibration equation that can be applied to data of unknown samples in order to predict their amounts of the constituents. Calculation of these equations requires a set of samples that span the expected range of amounts in the unknown samples. Calibration (known) and unknown samples have to be measured under exactly the same conditions using the same measurement technique. The measurements of the calibration samples, together with their amounts or concentrations, form the calibration data set from which the calibration equation (the model) is calculated. The model can then be applied to the unknown samples (validation data set) in order to predict their amounts or concentrations or to test the model’s validity (Duckworth, 1998).

4.3.1 Correlation analysis

The *correlation coefficient* (product-moment correlation coefficient or Pearson’s correlation coefficient) is a measurement of the strength of a relationship between two variables. It is defined as the covariance of the two variables divided by the product of their standard deviations; hence it may be thought of as a measure of standardised covariance. It is symbolized by r, and it ranges from 1.00 for perfect positive correlation, through zero for uncorrelated variables, to -1.00 for perfect negative correlation. Correlation coefficients may also be calculated for multiple regressions. Then it is an index of how well a dependent variable can be predicted from a linear combination of independent variables. It ranges from 0 (zero multiple correlation) to 1 (perfect multiple correlation) (Colman, 2001). When one variable is being used to predict another, the *coefficient of determination* ($R^2$) indicates the proportion of variance in one variable that is explained by the other (given certain linear assumptions), or the proportion of variance shared by the two variables: if $r = 0.50$, then $R^2 = 0.25$, indicating that 25 per cent of the variance in one of the variables is explained by the other, or that 25 per cent of the variance is shared by the two variables. It is measure of the goodness of fit of a set of scores to a linear model and can range in value from 0 to 1, with 1 indicating perfect fit (Colman, 2001).
4.3.2 Regression analysis

Bivariate linear regression

Bivariate linear regression (BLR) is the analysis of the relationship between a predictor variable (also called independent variable), such as a vegetation index and a dependent variable, e.g. a forest canopy attribute. The relationship is usually expressed as a linear model in which the value of the dependent variable is equal to a weighted value of the independent variable plus a constant:

\[ Y = b_0 + b_1 X \]

When more than one predictor variable is involved, it is called multiple regression.

Multiple linear regression

The spectrum of a sample measured by a spectrometer has many spectral data points that can be used to generate the calibration model. The problem is to find those wavebands that have a relationship to the concentration of the constituents of interest. Mostly multivariate models are used to find the best relation between the independent variables (spectral reflectances or absorbances) and the dependent data (concentrations or contents). A technique called least squares regression is employed to solve for the model equation; it calculates the coefficients of a given equation such that the differences between the known spectral responses and the predicted spectral responses are minimized. The classical least squares model (CLS) calculates reflectance at a single wavelength as an additive function of the constituent concentrations. The main drawback of CLS is that it can be only applied to systems in which the concentration of every constituent in the sample is known. The inverse least squares model (ILS) or multiple linear regression (MLR) assumes that the concentration of a constituent is a function of the reflectances at a series of given wavelengths. The advantage of the MLR is that it can accurately build models for complex mixtures when only some of the constituent concentrations are known; the only requirement is selecting wavelengths that correspond to the reflectances of the desired constituents (Duckworth, 1998).

Multiple linear regression (MLR) is a statistical method for analysing the joint and separate influences of two or more predictor variables on a dependent variable, usually through a linear model. The form of a multiple regression equation for \( k \) independent variables is a weighted sum:

\[ Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k \]

where \( Y \) is the estimated value of the dependent variable, \( b_0 \) is the \( Y \) intercept, \( X_1 \) is the value of the first independent variable, \( X_2 \) the value of the second independent variable, \( b_1 \) the slope associated with \( X_1 \), and \( b_2 \) the slope associated with \( X_2 \). The slope is a standardised regression
coefficient indicating the relative importance of the corresponding independent variable in determining the predicted value of the dependent variable (Colman, 2001). In this research and typically in spectroscopy or imaging spectrometry, spectra of a leaf or canopy target recorded at a large number of wavebands serve as the independent variables and are used to predict structural and chemical variables of that target.

**Stepwise multiple linear regression**

Stepwise multiple linear regression (SMR) is a technique for choosing wavebands to include into a regression model. An important underlying assumption is that only some, but not all wavebands do have an important explanatory effect on the leaf chemistry. Forward SMR starts with no variables in the regression equation; at each step, the most statistically significant variable is added, as specified by the entry criteria (highest F-value or lowest p-value); at the same time the procedure computes the removal statistic for each variable and removes variables if possible (as determined by the removal criteria). Stepping is terminated when neither entry nor removal of variables can be performed (StatSoft, Inc., 2002).

Statistical significance is a property of the results of an empirical investigation suggesting that they are not due to chance alone. The 5 per cent level of significance has become conventional in the empirical sciences; this means that results are normally considered to be statistically significant if statistical tests show that the probability of obtaining, by chance alone, results at least as extreme as those obtained is less than 5 percent, usually written $p < 0.05$. If the significance test indicates that the probability is sufficiently small, the investigator is justified in rejecting the null hypothesis and therefore also in accepting its logical negation, namely the alternative hypothesis. If the probability is not sufficiently small, then the correct interpretation is that there is no evidence of an effect, not that there is evidence of no effect (Colman, 2001).

SMR, although very useful to estimate leaf chemicals from foliar spectra, can be criticised in three points (Curran, 1989): I) Overfitting of wavebands in the calibration equation can occur when the number of samples is smaller than the number of wavebands. When the number of wavebands increases in the calibration model, the chance that unknown samples will vary in exactly the same way decreases, because the model starts to include the spectral noise unique to the training data set. Therefore, only those wavebands that have a causal relationship with the chemical of interest should be included in the calibration equation. II) Intercorrelation of chemicals may result in the selection of a waveband in the calibration model that is not related to a chemical bond of the constituent A but to the chemical bond of another constituent B that is highly correlated with A. III) It is difficult to explain why a particular waveband was included into the model while another waveband equally suited in terms of its relationship to chemical bonds was omitted.
Multiple regression is based on four main assumptions that can be easily violated when adapting the approach to remote sensing (Curran, 1989): I) The assumption that the relationships between reflectance and chemical concentration are near-linear is reasonable. II) The assumption that we can extract the vegetation spectra of interest from imaging spectrometry data is reasonable if continuous vegetation stands are studied. III) The assumption that the relationship between spectra and foliar chemicals is not disturbed by canopy geometry or phenology is reasonable if stands of a single species of vegetation are studied. IV) The assumption of accuracy in the measurements of chemical concentrations in the vegetation canopy is difficult to fulfil. It consists of two additional assumptions: i) the chemical composition of a canopy does not vary within a given volume; ii) the concentrations of a chemical measured in the laboratory does not differ from its concentration in the field due to a) transformation of the chemicals during the isolation process, b) the occurrence of many different forms of one single chemical, e.g. lignin, and c) the method chosen to measure the chemical.

4.3.3 Artificial neural networks

Neural networks are very sophisticated modeling techniques capable of modeling extremely complex, nonlinear functions. For many years linear modeling has been the commonly used technique for most modeling tasks, but where the linear approximation was not valid the models suffered accordingly.

Multilayer Perceptrons (MLP) (or multi layer feed forward networks) are perhaps the most popular network architecture in use today. In MLP, signals flow from inputs, forward through any hidden units and reach the output units. The units (or neurons) each perform a biased weighted sum of their inputs and pass this activation level through a transfer function to produce their output. The hidden and output layer neurons are each connected to all of the units in the preceding layer. The network has a simple interpretation as a form of input-output model, with the weights and biases the free parameters of the model. Such networks can model functions of almost arbitrary complexity, with the number of layers, and the number of units in each layer, determining the function complexity (Haykin, 1994; Bishop, 1995).

In supervised learning, the network learns to infer the relationship between the examples of inputs together with the corresponding outputs. Training data is usually taken from historical records or from simulations. The neural network is then trained using one of the supervised learning algorithms (e.g. back propagation), which uses the data to adjust the network's weights and thresholds in such a way that the error in its predictions on the training set is minimised. If the network is properly trained, it has then learned to model the (unknown) function that relates the input variables to the output variables, and can subsequently be used to make predictions where the output is not known (StatSoft, Inc., 2002).
The best-known example of a neural network training algorithm is back propagation (Rumelhart et. al., 1986). In back propagation, the gradient vector of the error surface is calculated. This vector points along the line of steepest descent from the current point. A sequence of moves in the direction of the vector will eventually find a minimum. The step size is proportional to the slope and to a special constant, the learning rate. The correct setting for the learning rate is application-dependent, and is typically chosen by experiment. The algorithm progresses iteratively, through a number of epochs. On each epoch, the training cases are each submitted in turn to the network, and target and actual outputs compared and the error calculated. This error, together with the error surface gradient, is used to adjust the weights, and then the process repeats. The initial network configuration is random and training stops when a given number of epochs elapses or when the error reaches an acceptable level (StatSoft, Inc., 2002).

The training data set includes a number of cases, each containing values for a range of input and output variables. Gathering cases and selecting input variables are critical issues of neural network design. The number of cases required for neural network training is related to the (unknown) complexity of the underlying function which the network is trying to model. For most practical applications, the number of cases required will be hundreds or thousands. As the number of variables increases, the number of cases required increases nonlinearly. (StatSoft, Inc., 2002). Problems of over-learning (or over-fitting) and selection of input variables are discussed in section 6.3.6.

**4.3.4 Ratio-based and orthogonal vegetation indices**

Broadband vegetation indices are based on discrete bands, usually located in red and near-IR wavelengths (Table 4.1). They are generally divided into ratio-based indices (RVI, NDVI, MVI, RSR) and orthogonal indices (PVI, SAVI, TSAVI). Ratio-based indices are calculated independently of soil reflectance properties, whereas orthogonal indices include parameters referring to the soil line. Ratio-based VIs tend to minimize the effects of illumination conditions while enhancing the contrast between vegetation and soil (Baret & Guyot, 1991). Orthogonal VIs aim to compensate for soil and background effects.

Most of the VIs are based on the reflectance in the red and near-IR part of the spectrum. With hyperspectral data, VIs can generally be calculated for all possible two-band combinations:

\[
RVI = \frac{\rho_{bd1}}{\rho_{bd2}}
\]

\[
PVI = \frac{\rho_{bd1} - a\rho_{bd2} - b}{\sqrt{1 + a^2}}
\]
It is then assumed that the soil line concept, originally defined for the red–near-IR feature space, can be transferred into other spectral domains (Thenkabail et al., 2000). Hence, it is supposed that soil lines exist between all wavebands.

4.3.5 The red edge inflection point

As the red edge inflection point (REIP) is defined as the point of maximum slope in the reflectance spectrum it may be determined using the first derivative of the reflectance. This approach will yield precise calculations of the REIP on data acquired by field-spectrometers which collect spectra with a spectral resolution of about 1 nm. However, in the case of imaging spectrometer data with a spectral resolution of usually about 10 to 20 nm, the subtle shifts of the REIP may not be recognized as the vegetation spectrum is sampled in a limited number of points. Therefore, a function has to be fitted through the measured data points of the reflectance spectrum. Several methods have been developed for interpolation in the

| Table 4.1: Broadband vegetation indices investigated in this study. \( \rho = \) reflectance, TM = Thematic Mapper |
|---|---|---|---|
| Name | Abbreviation | Equation | Reference |
| Ratio vegetation index | RVI | \( \frac{\rho_{TM4}}{\rho_{TM3}} \) | Pearson & Miller, 1972 |
| Normalised difference vegetation index | NDVI | \( \frac{\rho_{TM4} - \rho_{TM3}}{\rho_{TM4} + \rho_{TM3}} \) | Rouse et al., 1974 |
| Perpendicular vegetation index | PVI | \( \frac{\rho_{TM4} - a \rho_{TM3} - b}{\sqrt{1 + a^2}} \) \( a = 0.9, b = 0.1 \) | Richardson & Wiegand, 1977 |
| Transformed soil-adjusted vegetation index | TSAVI | \( \frac{a(\rho_{TM4} - a \rho_{TM3} - b)}{a \rho_{TM4} + \rho_{TM3} - ab} \) \( a = 0.9, b = 0.1 \) | Baret et al., 1989 |
| Mid-infrared vegetation index | MVI | \( \frac{\rho_{TM4}}{\rho_{TM5}} \) | Fassnacht et al., 1997 |
| Greeness vegetation index | GVI | \( -0.2848 \rho_{TM1} - 0.2435 \rho_{TM2} - 0.5436 \rho_{TM3} + 0.7243 \rho_{TM4} + 0.0840 \rho_{TM5} - 0.1800 \rho_{TM7} \) | Christ and Cicone, 1984 |
| Reduced simple ratio | RSR | \( RVI \cdot \left[ 1 - \frac{\rho_{TM5} - \min \rho_{TM5}}{\max \rho_{TM5} - \min \rho_{TM5}} \right] \) | Brown et al., 2000 |
wavelength region of the red edge. An approach suggested by Guyot et al. (1992) calculates in a first step the reflectance at the red edge inflection point by linear interpolation. More sophisticated approaches are the inverted Gaussian model (IGM) and the La Grange interpolation (LGI) technique.

**Inverted Gaussian model**

The inverted Gauss model (IGM) (Bonham-Carter, 1988) fits a Gaussian normal function to the reflectance at the red edge

\[ R_{est}(\lambda) = R_s - (R_s - R_0) \cdot e^{-\frac{(\lambda-\lambda_0)^2}{2\sigma^2}} \]

where \( R_{est} \) is the estimated reflectance at wavelength \( \lambda \), \( R_s \) is the reflectance maximum ("shoulder" reflectance), usually at approximately 780-800 nm; \( R_0 \) is the reflectance minimum, usually at about 670-690 nm; \( \lambda_0 \) is the wavelength of the reflectance minimum; \( \sigma \) is the Gaussian shape parameter with unit nm. The estimated \( REIP_{IGM} \) is then the midpoint on the ascending part of the modelled curve:

\[ REIP_{IGM} = \lambda_0 + \sigma \]

The function is fitted through the measured reflectance data points \( R_{mes}(\lambda) \) by adjusting the values of \( R_s, R_0, \lambda_0 \) and \( \sigma \) in such a way that the root mean square error (\( rmse \)) is minimized according to:

\[ RMSE = \sqrt{\frac{\sum_{i=1}^{N} (R_{mes}(\lambda_i) - R_{est}(\lambda_i))^2}{N}} \]

where \( N \) is the number of data points recorded in the \( N \) corresponding wavebands.

**Lagrangian model**

Another approach, the Lagrangian interpolation (LGI), has been proposed by Dawson & Curran (1998). The technique is applied to the approximate first-derivative of the reflectance spectrum, which is computed as follows:

\[ D_{mes}(\lambda_j + \Delta\lambda) = \frac{R(\lambda_j) - R(\lambda_{j-1})}{\lambda_j - \lambda_{j-1}} \]

where \( D_{mes}(\lambda_j + \Delta\lambda) \) is the measured first-derivative transformation at the midpoint with wavelength \( \lambda_j + \Delta\lambda \) between the wavebands \( j-1 \) and \( j \); \( R(\lambda_{j-1}) \) and \( R(\lambda_j) \) are the reflectances at the bands \( j-1 \) and \( j \), respectively. A second order polynomial is fitted directly to three bands of the first-order derivative spectrum:
\[ D_{\text{est}}(\lambda) = \frac{(\lambda - \lambda_i) \cdot (\lambda - \lambda_{i+1})}{(\lambda_{i-1} - \lambda_i) \cdot (\lambda_i - \lambda_{i+1})} \cdot D_{\text{mes}}(\lambda_{i+1}) + \frac{(\lambda - \lambda_{i-1}) \cdot (\lambda - \lambda_i)}{(\lambda_i - \lambda_{i-1}) \cdot (\lambda_i - \lambda_{i+1})} \cdot D_{\text{mes}}(\lambda_i) + \frac{(\lambda - \lambda_i) \cdot (\lambda - \lambda_{i+1})}{(\lambda_{i+1} - \lambda_i) \cdot (\lambda_i - \lambda_{i+1})} \cdot D_{\text{mes}}(\lambda_i) \quad 4.11 \]

where \( D_{\text{est}}(\lambda) \) is the first derivative estimated by the LGI model at any wavelength \( \lambda \); \( \lambda_i \) is the band having the maximum first derivative; \( \lambda_{i-1} \) and \( \lambda_{i+1} \) are the bands on the left and right side of \( \lambda_i \), respectively; \( D_{\text{mes}}(\lambda_i) \), \( D_{\text{mes}}(\lambda_{i-1}) \) and \( D_{\text{mes}}(\lambda_{i+1}) \) are the measured first derivative values.

The REIP is located at the wavelength \( REIP_{\text{LGI}} \) where \( D_{\text{est}}(\lambda) \) is maximum; to determine this position (and thus, the position of maximum slope of reflectance), the first derivation on \( D_{\text{est}}(\lambda) \) 4.11 is performed, representing the second derivative of the reflectance. The equation is resolved, for when the first derivative of \( D_{\text{est}}(\lambda) \) is zero, giving:

\[ REIP_{\text{LGI}} = \frac{A \cdot (\lambda_i + \lambda_{i+1}) + B \cdot (\lambda_{i-1} + \lambda_{i+1}) + C \cdot (\lambda_{i-1} + \lambda_i)}{2 \cdot (A + B + C)} \quad 4.12 \]

where

\[ A = \frac{D_{\text{mes}}(\lambda_{i-1})}{(\lambda_{i-1} - \lambda_i) \cdot (\lambda_i - \lambda_{i+1})}, \quad B = \frac{D_{\text{mes}}(\lambda_i)}{(\lambda_i - \lambda_{i-1}) \cdot (\lambda_i - \lambda_{i+1})}, \quad C = \frac{D_{\text{mes}}(\lambda_{i+1})}{(\lambda_{i+1} - \lambda_i) \cdot (\lambda_i - \lambda_{i+1})} \quad 4.13 \]

The characteristics of both methods for the calculation of the red edge inflection point were compared by Dawson & Curran (1998): In the case of the linear and the inverted Gaussian model, certain parameters such as \( R_s \) and \( R_0 \) have to be predefined. These parameters depend, apart from the leaf optical properties, on canopy optical properties which are unknown. The Lagrangian model, on the other hand, does not need \textit{a priori} knowledge of the spectrum and is therefore better suited to be used on canopy spectra than the other approaches. Furthermore, it forces the interpolation curve through the measured data points, thus taking into account the curvature of the function, whereas the IGM ignores the curvature as it fits a normal curve. The drawback of the Lagrangian method is the higher sensitivity to spectral noise as it is calculating the REIP from the derivative spectra.

### 4.3.6 Derivative spectra

The derivative of a spectrum is the rate of change of reflectance with respect to wavelength. The approximate first-order derivatives can be calculated according to

\[ D(\lambda_j + \Delta \lambda) = \frac{R(\lambda_j) - R(\lambda_{j-1})}{\lambda_j - \lambda_{j-1}} \quad 4.14 \]
where $D(\lambda_{j}+\Delta\lambda)$ is the first-derivative transformation at the midpoint with wavelength $\lambda_{j}+\Delta\lambda$ between the wavebands $j-1$ and $j$; $R(\lambda_{j-1})$ and $R(\lambda_{j})$ are the reflectances at the bands $j-1$ and $j$, respectively. Derivative spectra have been widely used in spectroscopy and remote sensing. Reflectance spectra may suffer from background signals and albedo effects. In contrast, derivatives were shown to be less sensitive to variations of soil background, illumination, and surface albedo (Demetriades-Shah et al., 1990). The main disadvantage of derivatives is their sensitivity to sensor noise. The decrease in the signal-to-noise ratio (SNR) with increasing derivative order often limits the usage of higher order derivatives.

4.3.7 Continuum-removal and band-depth normalisation

Absorptions in a spectrum are composed of the continuum and individual features, where the continuum is the background absorption onto which other absorption features are placed over. The continuum is simply an estimate of the other absorptions present in the spectrum, not including the one of interest (Clark & Roush, 1984). In practice, linear segments can be used to approximate the continuum line. Once this line is established, the continuum removed spectra ($CR$) are calculated by dividing the original reflectance values ($\rho_o$) by the values of the continuum line ($\rho_{cl}$) at the corresponding wavelength (Kokaly & Clark, 1999) as $CR=\rho_o/\rho_{cl}$. As non-linear multiplicative effects shape reflectance spectra, the continuum

![Figure 4.1: Continuum-removal for chlorophyll absorption in vegetation (from Clark, 1999).](image)
should be removed by division and only by subtraction when working with absorption coefficients (Clark, 1999). The removal of the continuum isolates spectral features and makes them comparable. Continuum removal, for instance, for the chlorophyll feature of green vegetation (Figure 4.1) may reveal detailed spectral variations of absorption that are not obvious in the original reflectance spectra (Clark, 1999).

Continuum-removed absorption features can be compared by scaling them to the same depth at the band centre. The band depth ($\rho_D$) at each wavelength in the absorption feature is computed from the continuum-removed reflectance ($\rho_{CR}$) as $\rho_D = 1 - \rho_{CR}$. The normalized band depths ($\rho_{DN}$) are calculated by dividing the band depth of each waveband ($\rho_D$) by the band-depth at the band centre ($\rho_{Dc}$) as $\rho_{DN} = \rho_D / \rho_{Dc}$ (Kokaly & Clark, 1999). The main purpose of continuum-removal and band-depth-normalization is the minimization of effects that extraneous factors may have on reflectance spectra to highlight shape differences of absorption features. Shape differences are then correlated with variations in foliar biochemistry through multiple stepwise linear regression.

4.3.8 Non-linear least-squares spectral matching

Gao & Goetz (1994, 1995) developed the non-linear least squares spectrum-matching technique to determine whether reflectance spectra of green leaves in the 1.0-2.5 μm region contain only liquid water absorption features or contain both the absorption features of liquid water and that of other leaf compounds. This method attempts to calculate a fresh leaf spectrum as a non-linear combination of a measured leaf water spectrum and a measured dry leaf spectrum. The leaf water spectrum was obtained by measuring reflectances of glass beads mixed with liquid water. The dry leaf spectrum was obtained by measuring reflectances of dried ground leaves. Approximate absorption coefficients ($k_\lambda$) of liquid water and of dry leaves were calculated from reflectance $R_\lambda$ as $k_\lambda = -\log(R_\lambda)$. The authors found that besides water, dry leaf components, such as lignin and cellulose, control the reflectance spectra of green vegetation (oak and cotton leaves) in that spectral domain.

The approach by Gao & Goetz (1994) was modified to estimate protein concentration in Norway spruce needles. Absorption coefficients of leaf water and protein were obtained from Dawson et al. (1998). The spectral reflectance of fresh needles $R_{Fmod}(\lambda)$ was modeled from the spectral absorption coefficients of liquid water $k_W(\lambda)$ and protein $k_P(\lambda)$ as

$$R_{Fmod}(\lambda) = (A + B\lambda)e^{-(C_Wk_W(\lambda) + C_Pk_P(\lambda))}$$

where $C_W$ and $C_P$ represent concentrations of water and protein, respectively, and the term $(A + B\lambda)$ represents the background level of the calculated spectrum. $A$, $B$, $C_W$ and $C_P$ were iteratively adjusted in such a way that the sum of the squared differences between the measured reflectance $R_{Fmes}(\lambda)$ and the modeled reflectance $R_{Fmod}(\lambda)$ for each given waveband

...
was minimized. The spectral reflectance of water $R_{\text{W mod}}(\lambda)$ was calculated by setting the parameter $C_P(\lambda)$ to zero in equation 4.15), giving:

$$R_{\text{W mod}}(\lambda) = (A + B\lambda)e^{-(C_P(\lambda))}$$

4.4 Physically-based reflectance modeling

4.4.1 Canopy reflectance models

Starting from the 1970’s a number of radiative transfer models, so-called canopy reflectance (CR) models have been developed. There are four broad categories of CR models (Goel, 1988; Roberts, 2001; Atzberger, 2003): Geometrical models (e.g., Li & Strahler, 1985) describe the vegetation canopy using opaque geometric shapes (cones or cylinders), which cast shadows on the ground; crown transparency is not considered. They are used to describe forests or shrublands, where shadowing plays an important role. Turbid medium or analytical models (e.g. Suits, 1972; Verhoef, 1984) describe the canopy as a horizontally homogenous and transparent layer that contain small vegetation elements which are treated as absorbing and scattering particles of a given geometry and density. These models are better suited to describe dense leafy canopies such as crops, but shadowing is not taken into account. Hybrid models are combinations of geometrical and turbid medium models. The geometric shapes representing, for instance, tree crowns, are not considered as being opaque but are treated as a turbid medium. This model type is often more complex but also closer to reality. FLIM

Figure 4.2: Schematic view of the link between LIBERTY, SAIL, and FLIM models showing the input variables used for the simulation. Modified after Zarco-Tejada et al. (2004a).
(Rosema et al., 1992) can be considered a very simple hybrid model, whereas INFORM (Atzberger, 2000) is of more sophisticated nature. Ray-tracing or computer simulation models (e.g., Gastellu-Etchegorry et al., 1996) compute the path of each photon through the canopy using random numbers to determine the direction of the incident rays, the location of the vegetation elements, the direction of scattering, and so on. Due to the complex structure and the large number of input parameters required, these models are difficult to invert (Goel, 1988; Roberts, 2001; Atzberger, 2003).

4.4.2 The Invertible Forest Reflectance Model

The Invertible Forest Reflectance Model (INFORM) (Atzberger, 2000) simulates the bi-directional reflectance of forest stands between 400 and 2500 nm. INFORM is essentially a combination of FLIM (Rosema et al., 1992), SAIL (Verhoef, 1984), and PROSPECT (Jacquemoud & Baret, 1990) or LIBERTY (Dawson et al., 1998). A schematic view of the link between the models is given in Figure 4.2.

The Forest Light Interaction Model (FLIM) takes into account both, the effects of shadowing and crown transmittance (Rosema et al., 1992). In this model, the forest is considered as a discontinuous canopy layer with tree crowns and gaps. The underlying assumption is that canopy reflectance is primarily controlled by coverage and crown transmittance (computed from LAI), and that background characteristics, spatial distribution of trees, and stand structure are of secondary importance (Gemmel & Varjo, 1999). FLIM calculates stand reflectance \( R \) in a given spectral band as

\[
R = R_C \cdot C + R_G \cdot G
\]

where \( R_C \) is the crown reflectance at infinite crown depth and \( R_G \) is the background reflectance. \( C \) is the “crown factor”:

\[
C = (1 - T_s \cdot T_o) \cdot c_s \cdot c_o
\]

where \( T_s \) is the average crown transmittance in sun direction and \( T_o \) is the average crown transmittance in observation direction; both \( T_s \) and \( T_o \) are exponentially related to LAI; \( c_o \) is the ground coverage by crowns in observation direction; \( c_s \) is the ground coverage by shadow in sun direction.

\( G \) is the “ground factor”:

\[
G = F_a \cdot T_s \cdot T_o + F_b \cdot T_s + F_c \cdot T_o + F_d
\]

where \( F_s \) are ground surface fractions; \( F_a \) represents tree crowns with shadowed background, \( F_b \) tree crowns with sunlit background, \( F_c \) shadowed open space, and \( F_d \) sunlit open space. The fractions in equation 4.19 are calculated from \( c_o \) and \( c_s \) and additionally take into account
the correlation between $c_o$ and $c_s$ which is in turn calculated from geometrical considerations using the tree height ($H$) and the crown diameter (Rosema et al., 1992).

The observed ground coverage by crowns $c_o$ under an observation zenith angle $\theta_o$ is given by

$$c_o = 1 - e^{-k \cdot SD \cdot \cos \theta_o}$$  \hspace{1cm} (4.20)

where $SD$ is the tree or stem density [ha$^{-1}$] and $k$ is a constant representing the average tree crown horizontal area [ha]. Similarly to the observed ground coverage by crowns, the ground coverage by shadow $c_s$ under a solar zenith angle $\theta_s$ is defined as

$$c_s = 1 - e^{-k \cdot SD \cdot \cos \theta_s}$$  \hspace{1cm} (4.21)

The constant $k$ representing the average horizontal area of a single tree crown [ha] can be derived from

$$k = \frac{\pi \cdot (0.5 \cdot CD)^2}{10000}$$  \hspace{1cm} (4.22)

where $CD$ is the crown diameter [m].

In FLIM, crown transmittance terms ($T_s$ and $T_o$) are assumed to be exponentially related to LAI; this neglects the effects of leaf geometry and leaf optical properties on the transmittance. Therefore in INFORM crown transmittance is calculated by the SAIL model (Verhoef, 1984). To compute crown transmittance for observation $T_o$ and sun direction $T_s$ (for a given leaf transmittance $\tau$ and leaf reflectance $\rho$), leaf area index (LAI), average leaf inclination angle (ALA), observation angle ($\theta_o$), sun angle ($\theta_s$), the relative azimuth angle between sun and sensor ($\psi$) and the fraction of diffuse radiation ($skyl$) have to be specified:

$$T_o = f_{SAIL}(LAI, ALA, \tau, \rho, \theta_o, \psi, skyl)$$  \hspace{1cm} (4.23)

$$T_s = f_{SAIL}(LAI, ALA, \tau, \rho, \theta_s, \psi, skyl)$$  \hspace{1cm} (4.24)

The original version of INFORM accounted for the reduction of transmittance by woody components within the canopy; the transmittance calculated by the SAIL model was reduced by a wood factor that represented the shadow cast on the ground by woody parts in a leafless crown (Atzberger, 2000). The correction for woody components was not applied in the present research as their influence is difficult to quantify.

In the FLIM model, the crown reflectance at infinite crown depth $R_C$ and the background reflectance $R_G$ (equation 1) are both derived empirically from the image data itself. INFORM, conversely, computes both terms using the SAIL model. For a given illumination and observation geometry, $R_G$ is derived as
where $LAI_U$ and $ALA_U$ are the LAI and the ALA, respectively, of the understorey vegetation and $\rho_{\text{soil}}$ is the soil reflectance. A soil reflectance spectrum must be provided from measurements. For a given illumination and observation geometry, $R_C$ is computed as

$$R_C = f_{\text{SAIL}}(LAI_{\text{inf}}, ALA, \tau, \rho, \rho_{\text{soil}}, \theta_o, \theta_s, \psi, \text{sky})$$

4.26

where $LAI_{\text{inf}}$ is the LAI at infinite crown depth.

In the initial version of INFORM (Atzberger, 2000), leaf spectral reflectance and transmittance were computed by the PROSPECT model (Jacquemoud & Baret, 1990) since only deciduous forests were considered. However, PROSPECT does not very well represent the optical properties of needle shaped leaves of conifer forest stands. Consequently, the LIBERTY model (Dawson et al, 1998) has been used as a replacement for PROSPECT. LIBERTY (Leaf Incorporating Biochemistry Exhibiting Reflectance and Transmittance Yields) is a general purpose radiative transfer model for simulating the reflectance and transmittance spectra of a conifer needle, or a stacked layer of needles between 400 and 2500 nm. It treats a needle as an aggregation of cells and calculates multiple scattering between cells. In LIBERTY, needle transmittance ($\tau$) and reflectance ($\rho$) are computed as a function of three structural parameters (cell diameter ($d$), intercellular air space ($i$), leaf thickness ($t$)) and the combined absorption coefficients of leaf chemicals (chlorophyll ($c_{AB}$), water ($c_{w}$), lignin and cellulose ($c_L$), and protein ($c_P$)):

$$\tau = f_{\text{LIBERTY}}(d, i, t, c_{AB}, c_{w}, c_L, c_P)$$

4.27

$$\rho = f_{\text{LIBERTY}}(d, i, t, c_{AB}, c_{w}, c_L, c_P)$$

4.28

In a summary, INFORM calculates forest reflectance $R$ as a function of internal canopy parameters, internal leaf parameters and external parameters:

$$R = f_{\text{INFORM}}(SD, CD, H, LAI, ALA, LAI_{\text{inf}}, LAI_U, ALA_U, \rho_{\text{soil}}, d, i, t, c_{AB}, c_{w}, c_L, c_P, \theta_o, \theta_s, \psi, \text{sky})$$

4.29

4.4.3 Inversion of reflectance models

In the forward model the CR model computes the spectral reflectance for a certain set of leaf or canopy parameters. Forward modeling provides an efficient way to investigate the effects of forest canopy attributes on reflectance signatures and may improve the understanding of how the canopy structure influences the radiation field. For the retrieval of those canopy variables from measured signals it is necessary to invert the model. Few studies exist, where CR models were inverted to estimate forest characteristics from remote sensing observations (section 2.3.3).
Due to the complex character of a CR model, an analytical solution is not possible and numerical approaches have been developed to successfully invert CR models. Traditional inversion of reflectance models employs an optimisation technique to estimate the model parameters by minimising a merit function (Goel, 1989). An iterative process is necessary to find the optimal estimates of these parameters. The main drawbacks of this method are i) difficulty in achieving globally optimal and stable results, ii) difficulty in retrieving more than 2 parameters simultaneously and iii) requirement of substantial computation time (Gong et al., 1999).

Recently, Artificial Neural Networks (ANN) have been employed for reflectance model inversion (Gong et al., 1999; Udelhoven et al., 2000; Kimes et al., 2002) in order to overcome the above mentioned limitations. To invert a reflectance model using an ANN the following steps have to be carried out (Gong et al., 1999): A) Determination of the structure of the ANN, in particular the number of nodes in the input, hidden, and output layer; each spectral band to be used corresponds to a node in the input layer; the number of parameters to be retrieved is the same as the number of nodes in the output layer. B) Simulation of the reflectance spectra from known parameter sets using the reflectance model. C) Presentation of known parameter sets and their corresponding reflectance spectra to the ANN; during the training step an algorithm is used to adjust weights and biases of the ANN in such a way that the root mean squared error (rmse) between estimated and true parameters is minimal. D) Input of measured reflectance spectra to the previously trained ANN and comparison of parameters estimated by the ANN to measured ones.

### 4.5 Validation

*Cross-validation* is a method of assessing the accuracy and validity of a calibration model. From the available data set one sample is selected and its spectrum and the corresponding attribute data are removed. The remaining training set samples are used to perform the calibration calculations. The calibration model is then used to predict the attribute of interest (e.g. concentration) of the removed sample. The previously left out sample is placed back into the training data set and a different sample is selected. Each sample is left out once and predicted by the training samples. Because the predicted samples are not the same as the samples used to build the model the cross-validated root mean squared error (rmse) is a good indicator of the accuracy of the model in predicting unknown samples (Duckworth, 1998). The rmse is the square root of the mean squared error; it is calculated from the differences between estimated values \( Y_{est} \) and measured \( Y_{mes} \) values and has the same units as the original data:
Another advantage of cross-validation is its ability to detect outliers. If the predicted concentration for a single sample is far off the measured concentration, the sample is probably an outlier (Duckworth, 1998).

*Validation set prediction* tests the predictive model on a validation data set. The validation set can come either from a new set of samples measured under the same conditions as the calibration samples or a large set of training data can be split into two groups: a set of calibration data and a set of validation data, where the split could be 80/20, 50/50, etc. depending on the total number of samples. The main advantage of this validation technique is the ability to test the model’s performance with a completely different and independent data set which is most important to determine the model’s long term stability (Duckworth, 1998).

\[
rmse = \sqrt{\frac{\sum_{i=1}^{n}(Y_{est} - Y_{meas})^2}{n}}
\]
Chapter 5 Mapping forest cover type through image classification

5.1 Introduction

Despite the large amount of spectral bands in data collected from imaging spectrometers, it has frequently been found, that the forest categorical variables considered do not have unique spectral signatures that allow forest classes to be distinguished without confusion. From this point of view it is still an open issue if hyperspectral data contain more information relevant to the classification of forest cover than broadband multispectral data. It is also not clear if a large spatial resolution of a few meters allow a better forest mapping compared to a pixel size of a few tens of meters typical to most earth observation satellite system, such as Landsat TM.

Forest types as defined by species composition and age class have spectral reflectance characteristics that may be strongly influenced by variations in the structural attributes of the forest stand (Franklin et al., 2003). For instance, variations in a stand structural attribute, such as stem density (the number of trees or stems per hectare) may have a greater effect on forest reflectance than tree species composition. Therefore, other mapped environmental variables that control forest vegetation distributions, such as terrain variables or geology, or structural attributes, can be combined with image data in the classification process to aid discrimination of forest type.

Classification should be applied to remotely sensed data of the appropriate spatial, spectral and temporal resolution for the required forest attribute information. Until recently, the choice of remote sensing data to map forests was limited to available operational sensor types. With airborne imaging spectrometers that were successfully operated in recent years, data with high spatial and high spectral resolution has become available. This data type offers the opportunity to systematically investigate the selection of optimal remote sensing data for forest mapping.

Based on the previous outline, the objectives of this study were to (i) compare the classification performance of four different remote sensing datasets of different spectral and spatial resolution (Figure 5.1) to map forest species and age classes and (ii) to investigate if additional information on stem density can improve the classification results.
5.2 Implementation of additional information

Information about forest parameters that have an influence on the signal recorded from the forest canopy by a sensor can support the estimation of other (unknown) parameters. The aim was to use spatial information on an important structural attribute, stem density (number of stems per hectare) as an ancillary layer in the classification process of the HyMap data in order to better distinguish between certain classes of forest cover type. Stem density information was derived from black/white aerial orthophotos by an approach described in Atzberger & Schlerf (2002). The procedure was based on thresholds of the digital numbers in the image data followed by morphological operations (erosion and dilation) and tree crown identification using a moving window. The method proofed to be capable of estimating stem density with an accuracy of approximately 63 trees per hectare for stem densities varying between 200 and 800 stems per hectare (Figure 5.2). Stem density was estimated for the entire area of the Idarwald test site. On a stand basis, the mean stem density and the standard deviation were calculated. The resulting image of the Idarwald test site (25 m x 25 m ground resolution) is shown in Figure 5.3a). From this image, per-field mean and standard deviation of the reference stem densities have been calculated (Figure 5.3b, c) and integrated into the FoGIS of Idarwald.
CHAPTER 5 MAPPING FOREST COVER TYPE

Figure 5.2: Estimated against field-measured stem density (cross-validated) at Idarwald test site. From Atzberger & Schlerf (2002).

Figure 5.3: Stem density (unit: ha\(^{-1}\)) at the Idarwald obtained from black/white orthophotos. (a) pixel resolution, (b) per-stand mean, (c) per-stand standard deviation. Stand borders are shown for clarity. From Atzberger & Schlerf (2002).
5.3 Image classification

Since the classification was restricted to conifers, deciduous forest was masked out using a forest mask, which had been generated through an unsupervised classification of the HyMap data using the isodata algorithm. All subsequent processing steps are restricted to coniferous forest.

5.3.1 Data preprocessing

From the available wavebands, usually a best minimum subset is selected for classification. This holds in particular for hyperspectral data as classification algorithms may suffer from redundancy (e.g. highly correlated wavebands). Two different types of spectral preprocessing were carried out, generation of multispectral data and compression of hyperspectral data. To generate synthetic, multispectral broadband data, the radiometrically and geometrically corrected HyMap image was spectrally resampled to Landsat TM wavebands using the appropriate sensor function. Hyperspectral data was compressed because an important requirement for a successful classification is the reduction of the spectral dimensions of the data without losing information. In the present study, the minimum noise fraction (MNF) transformation was used to achieve the data compression. The MNF transformation is similar to the Principal Component Analysis (PCA), but orders the data according to decreasing signal-to-noise ratio instead of decreasing variance (Green et al., 1988; Lee et al., 1990). The MNF transformation was used to partition the data space into two parts: one associated with large eigenvalues containing a large amount of information and a second with low eigenvalues and a large amount of noise. By using only the first portion of the data, the noise is separated from the data, thus reducing the number of bands and improving spectral processing results (Kruse, 1999). The first 19 MNF fractions were separated from the remaining fractions and used for further analysis.

Table 5.1: Training data used in the forest classification

<table>
<thead>
<tr>
<th>Forest class</th>
<th>Number of stands used for training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spruce, &gt;80 yrs</td>
<td>11</td>
</tr>
<tr>
<td>Spruce, 50-80 yrs</td>
<td>16</td>
</tr>
<tr>
<td>Spruce, 30-50 yrs</td>
<td>11</td>
</tr>
<tr>
<td>Spruce, 10-30 yrs</td>
<td>10</td>
</tr>
<tr>
<td>Douglas fir, 30-50 yrs</td>
<td>5</td>
</tr>
<tr>
<td>Douglas fir, 10-30 yrs</td>
<td>5</td>
</tr>
</tbody>
</table>
Spatial degradation of the high spatial resolution data (5 m) was done by calculating the mean pixel value within a 30 m x 30 m array for each waveband.

5.3.2 Training

Discrimination of forest types from classification of remote sensing imagery requires that each category (forest types) has unique reflectance properties that are different from the other categories (Franklin et al., 2003). While this concept of unique spectral signatures has been successfully applied in hyperspectral applications for mineral mapping, it has been less successful in forest applications due to the similar spectral response of vegetation canopies. Despite expected difficulties, it was attempted within this research to separate different forest types as defined by tree species and stand age.

A number of multivariate classification techniques (also called pattern recognition) have been successfully used in the remote sensing of forests. In supervised (opposed to unsupervised) classification, the predefined classes (e.g. forest types) are described from observations of the independent variables (reflectances and ancillary data), and then an algorithm is used to assign new observations (e.g. pixel of unknown forest type) to classes. This process assumes that the distribution of forest types to be mapped is known beforehand, at least for some part of the area of interest. This information is typically derived from field sampling or from ancillary data, such as a Geographic Information System.

Figure 5.4: Reflectance spectra of forest types considered in the classification process. Note: point measurements of reflectance at discrete waveband positions were connected with lines for clarity.
Table 5.2: Summary of accuracies

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Overall Accuracy</th>
<th>Kappa Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: HyMap 5m</td>
<td>81.4</td>
<td>0.78</td>
</tr>
<tr>
<td>B: HyMap 30m</td>
<td>95.0</td>
<td>0.94</td>
</tr>
<tr>
<td>C: TM 5 m</td>
<td>55.9</td>
<td>0.47</td>
</tr>
<tr>
<td>D: TM 30 m</td>
<td>69.5</td>
<td>0.64</td>
</tr>
<tr>
<td>E: HyMap 5 m + Stem density</td>
<td>83.5</td>
<td>0.81</td>
</tr>
</tbody>
</table>

The available FoGIS of Idarwald (section 3.4) was used to collect training data. The aim of the training stage is to collect a set of statistics that describe the spectral response pattern for each forest class. It was decided to consider only forest types in the classification that were represented by a minimum of three forest stands. Each of the six forest types that were considered represented a unique combination of a tree species (Norway spruce or Douglas fir) with an age class (Table 5.1). Other conifer tree species, such as fir and pine were very scarce in the study area and could not be considered in forest mapping. Age classes represent the stand development of Norway spruce that is observed at Idarwald test site and were defined by Womelsdorf, the local forest manager (personal communication, cited from Vohland, 1997). For Douglas fir, no definition of age classes was available and the lower and upper boundaries used to define age classes of Norway spruce were applied. The polygons in the FoGIS representing the forest stands were used to extract spectral signatures for each class. Within each forest type class, subclasses were defined. Each forest stand (polygon) that was available in the FoGIS was considered as a subclass. The generation of training data involved collecting 5-15 polygons per class, each polygon consisting of 100-1000 pixels. In total, 44 subclasses and 52,844 pixels were obtained. Training data were divided randomly into three sets of pixels: 10 percent were used for training (training pixels), 25 percent for validation of the classification result (validation pixels) and the remaining 65 percent pixels were not used. The training setup was designed this way to reflect the small amount of ground truth commonly available in reality. A visual interpretation of the collected spectra revealed considerable overlap of reflectance values for certain classes. Figure 5.4 shows the mean spectra of the considered forest types computed from all training pixels available.

5.3.3 Classification

Only Spectral Angle Mapper (SAM) algorithm (section 4.2.2) was employed for supervised classification as the focus of the study was on the information content of various types of data rather than the performance of different classifiers. After the classification, the 44 subclasses obtained were condensed into the 6 main classes. Then post-classification smoothing was performed to remove single isolated image pixels (sieve and clump).
5.3.4 Validation

To assess the accuracy the validation pixels were used. A confusion matrix was generated from the validation pixels for each classification. Two measures of classification accuracy are reported. Overall accuracy (OAA) quantifies the percentage of cases correctly classified:

\[
OAA = \frac{\sum_{k=1}^{q} n_{kk}}{n} \times 100
\]

where \( n_{kk} \) is the number of correctly classified validation pixels (confusion matrix diagonals), \( q \) is the number of classes, and \( n \) is the total number of validation pixels used. The kappa coefficient accommodates for the effects of chance agreement:

\[
KAPPA = \frac{n\sum_{k=1}^{q} n_{kk} - \sum_{k=1}^{q} n_{k+} n_{+k}}{n^2 - \sum_{k=1}^{q} n_{k+} n_{+k}}
\]

where \( n_{k+} \) is the sum of the validation pixels in a class and \( n_{+k} \) is the sum of the classified pixels in that class.

5.4 Results and discussion

Table 5.2 lists overall accuracies and kappa coefficients obtained from classifying each of the 5 sets of data (2 types of spectral resolution x 2 types of spatial resolution plus 1 type high spectral and high spatial resolution combined with stem density). Table 5.3 lists detailed confusion matrices for all datasets.

The overall accuracy (OAA) and Kappa coefficient using the HyMap datasets (A and C) were considerably larger than using the TM datasets (B and D). These results confirm that hyperspectral data contain more information relevant to the mapping of forest type than broadband multispectral data. Contrary to broadband data analysis techniques, hyperspectral data offer the use of full-spectrum analysis techniques that exploit all the inherent information. As an example, it has already been shown that the estimation of canopy LAI using hyperspectral imagery can be achieved with accuracy greater than with a broadband sensor (Gong et al, 1992, see also Chapter 6).
The OAA using the dataset A (HyMap, 5 m) was 81 percent. This increased to 95 percent using HyMap degraded to 30 m spatial resolution (dataset B). Essentially the same tendency was obtained in classifying TM 5 m (dataset C) and TM 30 m (dataset D). Obviously, an increase in image variance due to higher spatial resolution led to problems in the classification process. This finding is supported by classification results obtained by Martin & Howarth (1989) who found that an increase in spatial resolution did not result in higher classification accuracies. In the case of forest cover, each tree crown which is distinguishable in the high resolution imagery will have a sunlit and shaded portion with a highly variable spectral response. Moreover, relatively small gaps in the forest cover where understorey vegetation or bare soil become visible might be resolved in the image. Presumably, these effects are largely smoothed out in the low resolution dataset with the effect of a lower spatial variance (Leckie

Table 5.3: Confusion matrices for 5 different datasets: A: hyperspectral, 5 m; B: multispectral, 5 m; C: hyperspectral, 30 m; D: multispectral 30 m, E: hyperspectral 5 m and orthophoto derived stem density; Spr. = Norway spruce, Dou. = Douglas fir; numbers indicate stand age.

<table>
<thead>
<tr>
<th>Class</th>
<th>Spr. &gt;80</th>
<th>Spr. 50-80</th>
<th>Spr. 30-50</th>
<th>Spr. 10-30</th>
<th>Dou. 30-50</th>
<th>Dou. 10-30</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>4.2</td>
<td>4.1</td>
<td>8.1</td>
<td>1.2</td>
<td>2.8</td>
<td>1.7</td>
<td>4.4</td>
</tr>
<tr>
<td>Spr. &gt;80</td>
<td>84.1</td>
<td>6.8</td>
<td>3.8</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
<td>16.4</td>
</tr>
<tr>
<td>Spr. 50-80</td>
<td>10.3</td>
<td>86.7</td>
<td>29.2</td>
<td>0.0</td>
<td>3.8</td>
<td>0.2</td>
<td>32.6</td>
</tr>
<tr>
<td>Spr. 30-50</td>
<td>1.1</td>
<td>2.2</td>
<td>58.1</td>
<td>0.7</td>
<td>2.0</td>
<td>0.4</td>
<td>15.1</td>
</tr>
<tr>
<td>Spr. 10-30</td>
<td>0.3</td>
<td>0.2</td>
<td>0.6</td>
<td>97.9</td>
<td>1.0</td>
<td>2.4</td>
<td>9.1</td>
</tr>
<tr>
<td>Dou. 30-50</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>89.5</td>
<td>0.8</td>
<td>14.5</td>
</tr>
<tr>
<td>Dou. 10-30</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.7</td>
<td>94.5</td>
<td>7.8</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Spr. &gt;80</th>
<th>Spr. 50-80</th>
<th>Spr. 30-50</th>
<th>Spr. 10-30</th>
<th>Dou. 30-50</th>
<th>Dou. 10-30</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>2.8</td>
<td>3.9</td>
<td>1.3</td>
<td>3.0</td>
<td>1.1</td>
<td>2.2</td>
<td>2.4</td>
</tr>
<tr>
<td>Spr. &gt;80</td>
<td>95.4</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>1.8</td>
<td>17.7</td>
</tr>
<tr>
<td>Spr. 50-80</td>
<td>1.6</td>
<td>92.3</td>
<td>4.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>25.5</td>
</tr>
<tr>
<td>Spr. 30-50</td>
<td>0.2</td>
<td>3.7</td>
<td>94.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>22.6</td>
</tr>
<tr>
<td>Spr. 10-30</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>97.0</td>
<td>0.0</td>
<td>0.1</td>
<td>6.2</td>
</tr>
<tr>
<td>Dou. 30-50</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>98.8</td>
<td>0.0</td>
<td>15.9</td>
</tr>
<tr>
<td>Dou. 10-30</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>95.9</td>
<td>9.6</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Spr. &gt;80</th>
<th>Spr. 50-80</th>
<th>Spr. 30-50</th>
<th>Spr. 10-30</th>
<th>Dou. 30-50</th>
<th>Dou. 10-30</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>3.7</td>
<td>1.8</td>
<td>7.0</td>
<td>0.6</td>
<td>20.0</td>
<td>4.4</td>
<td>6.5</td>
</tr>
<tr>
<td>Spr. &gt;80</td>
<td>32.1</td>
<td>15.1</td>
<td>2.3</td>
<td>0.0</td>
<td>0.5</td>
<td>0.0</td>
<td>10.4</td>
</tr>
<tr>
<td>Spr. 50-80</td>
<td>63.3</td>
<td>81.6</td>
<td>48.3</td>
<td>0.0</td>
<td>22.0</td>
<td>0.4</td>
<td>47.6</td>
</tr>
<tr>
<td>Spr. 30-50</td>
<td>0.0</td>
<td>0.1</td>
<td>38.3</td>
<td>14.3</td>
<td>3.8</td>
<td>0.6</td>
<td>10.4</td>
</tr>
<tr>
<td>Spr. 10-30</td>
<td>1.0</td>
<td>1.3</td>
<td>4.0</td>
<td>84.2</td>
<td>0.6</td>
<td>4.9</td>
<td>7.4</td>
</tr>
<tr>
<td>Dou. 30-50</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>46.7</td>
<td>20.3</td>
<td>9.5</td>
</tr>
<tr>
<td>Dou. 10-30</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>1.0</td>
<td>6.4</td>
<td>69.4</td>
<td>8.1</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
1990; Schlerf et al., 2003). Foody (2002) claimed that accuracies larger than 85 percent are required from remote sensing image classification. The requirement is not fulfilled by most of the datasets except the hyperspectral data low spatial resolution data which can be explained by the fact, that the classification problem encountered is not trivial. Forest classes, such as clearings have spectral reflectance characteristics that differ very much from the closed canopy stands considered in the study and are relatively easy to discriminate from the other forest types. Including such forest classes probably would have increased the accuracy of the classification.

The type of post-classification smoothing applied to the classified image considerably influenced the classification accuracy. Also the design of the validation procedure controlled the classification results to a certain amount. Buddenbaum (2004) used a modified validation approach. In this approach the forest stands were divided into spatially separated training and validation areas. The received accuracies were somewhat lower.

The OAA using the full set of HyMap 5 m data and stem density information (dataset E) was 83 percent and thus 2 percent larger than the OAA using HyMap 5 m (dataset A) alone. This result confirms that the classification of forest cover may be improved through addition of orthophoto derived stem densities into the classification process. However, the increase in classification accuracy was rather limited. Various other types of ancillary information have already been utilized in multispectral classification of vegetation cover: Wulder (1998), for instance, used texture and Franklin (1994) used digital elevation models as additional

<table>
<thead>
<tr>
<th>E) Ground Truth (Percent)</th>
<th>Class</th>
<th>Spr. &gt;80</th>
<th>Spr. 50-80</th>
<th>Spr. 30-50</th>
<th>Spr. 10-30</th>
<th>Dou. 30-50</th>
<th>Dou. 10-30</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>3.5</td>
<td>3.6</td>
<td>5.5</td>
<td>0.5</td>
<td>2.1</td>
<td>1.5</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>Spr. &gt;80</td>
<td>90.0</td>
<td>10.3</td>
<td>4.1</td>
<td>0.0</td>
<td>0.4</td>
<td>0.0</td>
<td>18.4</td>
<td></td>
</tr>
<tr>
<td>Spr. 50-80</td>
<td>5.6</td>
<td>84.3</td>
<td>25.4</td>
<td>0.1</td>
<td>3.4</td>
<td>0.1</td>
<td>30.2</td>
<td></td>
</tr>
<tr>
<td>Spr. 30-50</td>
<td>0.6</td>
<td>1.7</td>
<td>64.3</td>
<td>0.1</td>
<td>2.8</td>
<td>0.4</td>
<td>16.5</td>
<td></td>
</tr>
<tr>
<td>Spr. 10-30</td>
<td>0.4</td>
<td>0.1</td>
<td>0.6</td>
<td>99.3</td>
<td>0.5</td>
<td>2.4</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>Dou. 30-50</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>89.9</td>
<td>0.7</td>
<td>14.6</td>
<td></td>
</tr>
<tr>
<td>Dou. 10-30</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.9</td>
<td>94.9</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D) Ground Truth (Percent)</th>
<th>Class</th>
<th>Spr. &gt;80</th>
<th>Spr. 50-80</th>
<th>Spr. 30-50</th>
<th>Spr. 10-30</th>
<th>Dou. 30-50</th>
<th>Dou. 10-30</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>8.0</td>
<td>6.4</td>
<td>10.8</td>
<td>17.8</td>
<td>4.8</td>
<td>3.3</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Spr. &gt;80</td>
<td>44.1</td>
<td>10.6</td>
<td>1.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.5</td>
<td>11.2</td>
<td></td>
</tr>
<tr>
<td>Spr. 50-80</td>
<td>41.9</td>
<td>83.0</td>
<td>41.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>38.9</td>
<td></td>
</tr>
<tr>
<td>Spr. 30-50</td>
<td>6.0</td>
<td>0.0</td>
<td>45.7</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td>Spr. 10-30</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>77.2</td>
<td>0.2</td>
<td>3.2</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>Dou. 30-50</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>94.3</td>
<td>2.3</td>
<td>15.4</td>
<td></td>
</tr>
<tr>
<td>Dou. 10-30</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.7</td>
<td>90.8</td>
<td>9.2</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
information and observed an increase of overall classification accuracy compared to using remote sensing data on a pixel basis alone. Buddenbaum (2004) computed texture parameters from hyperspectral image data by means of geostatistics and used both, hyperspectral wavebands and texture channels together for mapping conifer forest. The results pointed towards an increase in classification accuracy.

In contrast to difficulties reported in the literature (Niemann, 1995; Vohland, 1997; Köhl & Lautner, 2001), discrimination of medium aged to old stands was feasible using the outlined classification strategy. Identification of old forest stands that hold large amounts of biomass is of particular interest to forest authorities to determine the time of harvest. The resulting forest type map is presented in Figure 5.5.
CHAPTER 5 MAPPING FOREST COVER TYPE

Figure 5.5: Map of forest type at Idarwald.
Chapter 6  Determination of forest structural attributes

6.1 Introduction

Estimation of structural and chemical forest characteristics from remotely sensed data can be generally achieved by empirical or physical approaches (Asner et al., 2003; Chapter 4). Empirical models relate a measure of reflectance (e.g., a vegetation index, VI) to the variable of interest through a regression equation. In a physically based approach, a forest reflectance model is inverted to estimate the structural characteristics of a forest canopy. Both approaches do have certain advantages and disadvantages. For instance, reflectance models can take into account effects of shadowing and multiple scattering while those effects represent major challenges to vegetation indices. On the other hand, empirical models though they need costly field work data are easy to apply.

The main objective of this chapter was to compare empirical and physically-based approaches in terms of their capability to determine forest structural attributes from hyperspectral remote sensing imagery. Another commonly used concept, spectral unmixing, was not implemented within this study, as spectral mixture models can be considered as simple reflectance models which are better suited for sparse vegetation canopies than for dense forests.

Table 6.1: Resulting subsets after stratification according to forest stand age

<table>
<thead>
<tr>
<th>Subset</th>
<th>Designation</th>
<th>Abbreviation</th>
<th>Age</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>total (pooled dataset)</td>
<td>t</td>
<td>10-148</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>medium to old</td>
<td>mo</td>
<td>30-148</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>old</td>
<td>o</td>
<td>80-148</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>medium</td>
<td>m</td>
<td>30-79</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>young</td>
<td>y</td>
<td>10-29</td>
<td>5</td>
</tr>
</tbody>
</table>
6.2 Developing predictive models

6.2.1 Introduction

The overall aim of this part of the study was to develop linear predictive models between vegetation indices and forest structural variables. More specific objectives were (i) to identify useful stand variables out of the measured ones or generate new stand variables that could be correlated with reflectance measured by the HyMap sensor, (ii) to determine spectral vegetation indices that are best suited for characterising those stand variables, and (iii) to compare and contrast traditional broad-band and hyperspectral VIs in terms of basic statistical characteristics of the predicted stand variables relative to the observed stand variables.

To improve LAI mapping it has been suggested to derive land cover specific VI-LAI

<table>
<thead>
<tr>
<th>Stand variable</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Range</th>
<th>Coefficient of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perimeter-at-breast height (PBH) [m]</td>
<td>1.09</td>
<td>0.32</td>
<td>1.63</td>
<td>0.27</td>
<td>1.36</td>
<td>0.294</td>
</tr>
<tr>
<td>Stem density (SD) [ha⁻¹]</td>
<td>640</td>
<td>458</td>
<td>2444</td>
<td>244</td>
<td>2200</td>
<td>0.716</td>
</tr>
<tr>
<td>Canopy cover (CO) [%]</td>
<td>48</td>
<td>11</td>
<td>70</td>
<td>30</td>
<td>40</td>
<td>0.229</td>
</tr>
<tr>
<td>Canopy height (H) [m]</td>
<td>25.1</td>
<td>7.9</td>
<td>39.1</td>
<td>5.0</td>
<td>34.1</td>
<td>0.315</td>
</tr>
<tr>
<td>Leaf area index (LAI) [m² m⁻²]</td>
<td>3.24</td>
<td>0.97</td>
<td>5.47</td>
<td>1.66</td>
<td>3.81</td>
<td>0.299</td>
</tr>
<tr>
<td>Stem biomass (SBM) [m³ ha⁻¹]</td>
<td>201</td>
<td>109</td>
<td>654</td>
<td>11</td>
<td>643</td>
<td>0.540</td>
</tr>
</tbody>
</table>

Table 6.2: Summary statistics for forest stand attributes (n=40)

<table>
<thead>
<tr>
<th>Stand variable</th>
<th>PBH</th>
<th>SD</th>
<th>CO</th>
<th>H</th>
<th>LAI</th>
<th>SBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBH</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>-0.87**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>-0.55**</td>
<td>0.72**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>0.88**</td>
<td>-0.86**</td>
<td>-0.58**</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAI</td>
<td>-0.72**</td>
<td>0.70**</td>
<td>0.68**</td>
<td>-0.70**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>SBM</td>
<td>0.81**</td>
<td>-0.59**</td>
<td>-0.20</td>
<td>0.80**</td>
<td>-0.48**</td>
<td>1.00</td>
</tr>
<tr>
<td>AGE</td>
<td>0.85**</td>
<td>-0.72**</td>
<td>-0.49**</td>
<td>0.82**</td>
<td>-0.66**</td>
<td>0.77**</td>
</tr>
</tbody>
</table>

Table 6.3: Linear correlation between forest stand variables (n=40); ** correlation coefficient significant at P<0.01.
relationships using land cover maps (Franklin et al., 1997; Chen et al., 2002; Cohen et al., 2003). This statistical stratification process is usually based on the species type. As in the present study just one single species is considered, it was decided to stratify the dataset according to stand age (Table 6.1). Information about stand age was provided by a Forest Geographic Information System (FoGIS) that had been compiled for the Idarwald (section 3.4).

6.2.2 Computation of vegetation indices

Both, hyperspectral and broadband VIs (section 4.3.4) were computed from the radiometrically and geometrically corrected HyMap spectra after reflectance spectra had been extracted from the HyMap imagery (section 3.3). Ratio-based and orthogonal broadband VIs (Table 4.1) were calculated after the HyMap data had been resampled to Landsat TM spectral bands involving the appropriate Landsat TM filter functions; the ratio-based VIs were RVI and NDVI; the orthogonal VIs were PVI and TSAVI; additional VIs were MVI and GVI. Narrowband RVI and PVI (section 4.3.4) were systematically calculated for all possible $114 \times 114 = 12,996$ band combinations. PVI requires site-specific soil line slopes (a) and intercepts (b). As no soil spectral data was available, standard values (a=0.9; b=0.1) were used. Additionally, the REIP was calculated using two different methods, the IGM and the LGI (section 4.3.5). Bivariate linear regression (section 4.3.2) was employed to evaluate the relationships between structural stand attributes and VIs. A cross-validation (leave one out) procedure (section 4.5) was used to validate the regression models.

6.2.3 Analysis of forest stand variables

From the summary statistics (Table 6.2) it can be seen that the attributes perimeter-at-breast height (PBH), canopy height (H), and leaf area index (LAI) have a similar variability, whereas canopy cover (CO) is less variable and stem density (SD) and stem biomass (SBM) are more variable. The large variability of SD can be attributed to the extraordinary large density values of young forest stands.

Table 6.3 lists linear correlation coefficients between forest stand attributes. Strong relationships between stand attributes were found for perimeter-at-breast height, stem density, and canopy height. These variables are closely related to stand age; with increasing age trees grow in height and diameter and perimeter-at-breast height and canopy height increase, while stem density is reduced due to thinning. An inverse relationship can be observed between LAI and age. The inverse relation between LAI and stand age seems to be related to the fact that during stand development, crowns of individual trees expand and increase utilization of available growing space. The point at which crowns of different trees begin to interact is considered as being the peak LAI after which a rapid decrease takes place due to competition.
between the individual trees (Vose et al., 1994). It has been shown that stands of slow growing species (Pinus contora) reached its maximum LAI at age 40 and that it lasted for about 30 years (Long & Smith, 1992; cited from Vose et al., 1994).

Prevalent forestry practices in Germany aim to maximise the profitability of a site. Thinning, the extraction of some of the young trees in a forest so that the remainder grow and develop fully, has an additional, major influence on the temporal LAI dynamic. At Idarwald, after a maximum LAI value at an age of 20-30 years is reached, stands of Norway spruce are usually thinned at an age of 30-50. Therefore, relatively low LAI values can be expected. After 60-70 years an increase in LAI may take place up to an age of 100 where logging can lead to gaps in

Table 6.4: Linear correlation between forest stand attributes and HyMap reflectance resampled to TM spectral bands (n=40); ** correlation coefficient significant at P<0.01, * correlation coefficient significant at P<0.05

<table>
<thead>
<tr>
<th>Stand variable</th>
<th>TM3</th>
<th>TM4</th>
<th>TM5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perimeter-at-breast height</td>
<td>-0.60**</td>
<td>-0.80**</td>
<td>-0.55**</td>
</tr>
<tr>
<td>Stem density</td>
<td>0.72**</td>
<td>0.86**</td>
<td>0.66**</td>
</tr>
<tr>
<td>Canopy cover</td>
<td>0.41**</td>
<td>0.69**</td>
<td>0.38*</td>
</tr>
<tr>
<td>Canopy height</td>
<td>-0.57**</td>
<td>-0.84**</td>
<td>-0.57**</td>
</tr>
<tr>
<td>Leaf area index</td>
<td>0.36*</td>
<td>0.76**</td>
<td>0.33*</td>
</tr>
<tr>
<td>Stem biomass</td>
<td>-0.50**</td>
<td>-0.64**</td>
<td>-0.51**</td>
</tr>
</tbody>
</table>
the canopy cover. Depending on the application of thinning measures, maximum variation of LAI can be expected at age 30-70 (personal communication Wagner, Womelsdorf; cited from Vohland, 1997).

In Figure 6.1, LAI is plotted against stand age for the probed Norway spruce stands of Idarwald test site. Here, a peak LAI can be assumed at an age of 20. From an age of 20 onwards, a gradual decline in LAI from up to age 150 is evident. Maximum variation of LAI occurs at age 60-70. Obviously, some stands have been thinned lately whereas others have been thinned long time ago. After an increase of LAI up to age 100 a decrease of LAI is caused by gaps related to logging measures. It can be concluded that the observed age course of Norway spruce based on the probed sample is a result of both natural circumstances and actual management practices. Concurrently, regeneration of leaf area index subsequent to thinning measures is also determined by the water and nutrient supply of a particular site. It is also clear from Figure 6.1, that LAI can vary considerably within a single age class and thus, information on LAI can not simply be derived from an age class map.

At first, relationships between the forest stand attributes and single band reflectance were examined (Table 6.4). The stand attributes were most strongly correlated with TM4 and less correlated with TM3 and TM5. Correlations were positive for stem density (SD), canopy cover (CO), and LAI and negative for perimeter-at-breast height (PBH), canopy height (H) and stem biomass (SBM). PBH, SD, and H were more strongly correlated with reflectance whereas CO, LAI, and SBM were less correlated with reflectance. Attributes related to canopy cover (LAI, CO) are usually expected to have negative correlations with red reflectance. At Idarwald, we found weak positive correlations between LAI and red reflectance (TM3). However, red reflectance varied only between 2.7 and 3.0 percent and possibly, saturation in the red has already occurred at low LAI values. Another explanation could be, that green leaves and the underlying soil and litter have similar reflectances in the red and that an increase in LAI or CO would not have an influence on red reflectance. Finally, effects of shading could be responsible that old stands of relatively low LAI also have low reflectances in the red domain. One could also expect negative correlations between LAI and reflectance in the mid-IR (TM5) (Brown et al., 2000). Although we found such negative correlations for wavebands located in the water absorption features, we observed positive correlations for wavebands located at the reflectance maximum between the 1.4 µm and 1.9 µm water absorption features. Again, shading could be a possible explanation for these findings.
The correlation coefficients (r) between structural attributes and both reflectance and first derivative reflectance spectra are presented as correlograms (Figure 6.2). The strength of the relation generally decreased from near infrared (near-IR) to mid infrared (mid-IR) wavebands and was greater for first derivative reflectance spectra in opposition to reflectance spectra. The decreasing r values with increasing wavelength from the near-IR to the mid-IR reflected the decrease in signal-to-noise ratio of the data. In this respect the correlograms were similar to those reported in other studies (e.g., Curran et al., 2001). Correlations were positive throughout all wavelength regions for stem density, canopy cover, and LAI and negative for stem perimeter (PBH), canopy height, and stem biomass. The near-IR region of reflectance spectra revealed strongest correlations followed by the green peak, the mid-IR region, and the chlorophyll absorption features in the red and blue wavebands. For the correlation between

Figure 6.2: Correlograms for six stand attributes and both reflectances (left) and first derivative reflectance spectra (right) (n=40). Data points in first derivative reflectance correlograms are plotted as lines for clarity.
reflectance and LAI, similar curves were obtained in other studies (Thenkabail et al., 2000). However, one major difference was the positive correlation between LAI and reflectance in the red which has been already discussed in the previous paragraph.

From the strong intercorrelations between the stand attributes (Table 6.3) it is evident that no single variable could be regarded as causal to the reflectance in the TM bands. Therefore, the principal components (PC) were computed on the forest stand variables in order to reduce the data and generate new variables that could be correlated with reflectance. The input variables were standardized to zero mean and unit standard deviation. From the five original variables (all variables except SBM) five PC were derived. The first 3 PC explained more than 95 percent of the variability of the original dataset (not shown). The eigenvectors are vectors which define the principal component axes in terms of the original coordinate axes (Richards, 1993). The eigenvectors of PC1 revealed moderately large positive loadings on PBH and H and a moderately large negative loading on SD (not shown). While principal components analysis is often preferred as a method for data reduction, principal factors analysis is often chosen when the goal of the analysis is to detect structure. In principal components analysis we assume that all variability in a stand variable should be used in the analysis. In principal factors analysis we only use the variability in a stand variable that it has in common with the other stand variables. The goal of the various existing rotational strategies is to obtain a clear pattern of loadings, explicitly, factors that are clearly marked by high loadings for some variables and low loadings for others (Stat Soft, Inc., 2004; Harman, 1976).

Table 6.5: Statistics of the first component factor of the varimax rotation (CF1) compared to the Eigenvector of the first principal component PC1 (n=40)

<table>
<thead>
<tr>
<th>Stand variable</th>
<th>Specific variance CF1</th>
<th>Loading CF1</th>
<th>Loading PC1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perimeter-at-breast height</td>
<td>0.1268</td>
<td>-0.9345</td>
<td>0.4638</td>
</tr>
<tr>
<td>Stem density</td>
<td>0.1275</td>
<td>0.9341</td>
<td>-0.4757</td>
</tr>
<tr>
<td>Canopy cover</td>
<td>0.5400</td>
<td>0.6782</td>
<td>-0.3972</td>
</tr>
<tr>
<td>Canopy height</td>
<td>0.1417</td>
<td>-0.9265</td>
<td>0.4633</td>
</tr>
<tr>
<td>Leaf area index</td>
<td>0.4107</td>
<td>0.7677</td>
<td>-0.4315</td>
</tr>
</tbody>
</table>

Table 6.6: Linear correlation between principal components and HyMap reflectance resampled to TM spectral bands (n=40); ** correlation coefficient significant at P<0.01

<table>
<thead>
<tr>
<th>Component</th>
<th>TM3</th>
<th>TM4</th>
<th>TM5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF1</td>
<td>0.65**</td>
<td>0.88**</td>
<td>0.61**</td>
</tr>
<tr>
<td>PC1</td>
<td>-0.61**</td>
<td>-0.89**</td>
<td>-0.57**</td>
</tr>
<tr>
<td>PC2</td>
<td>-0.18</td>
<td>-0.02</td>
<td>-0.19</td>
</tr>
<tr>
<td>PC3</td>
<td>-0.29</td>
<td>-0.02</td>
<td>-0.28</td>
</tr>
</tbody>
</table>
The factor analysis was computed using the varimax rotation. Again, the input variables were standardized to zero mean and unit variance. After standardisation, one common factor (CF)
was computed from the five original stand attributes (Table 6.5). The p-value of 0.08 failed to reject the null hypothesis of one common factor suggesting that this model provides a satisfactory explanation of the covariation in these data. A specific variance of 1 would indicate that there is no common factor component in that attribute while a specific variance of 0 would indicate that the attribute is entirely determined by common factors. The estimated specific variances indicated that the stand attributes PBH, SD, and H are determined by the common factor. However, the common factor did not clearly represent the attributes LAI and CO. For interpretation of the common factor, its loading has to be examined (Table 6.5). CF1 has very high loadings on the variables PBH, SD, and H, but only moderately high loadings on LAI and CO. The loadings obtained by CF1 reveal a much clearer pattern than those obtained by PC1. According to the loadings of CF1, old stands consisting of large trees at low density would have low scores on CF1. It was obvious, to interpret CF1 as a variable that is inversely related to stand age. The correlation between CF1 and stem biomass was moderately strong \( r = -0.74 \). Following an interpretation by Danson & Curran (1993), CF1 was regarded as a variable inversely related to crown volume (VOL).

In a next step, the common factor and the principal components were correlated with the reflectance data in the TM bands (Table 6.6). Statistically significant positive relationships were found between VOL and TM reflectance. This indicated, that stands for which a high reflectance was recorded had a high score on CF1 and thus, a low canopy volume. Also PC1 showed a close relationship to TM reflectance. No correlations were observed between PC of higher order and reflectance. From Figure 6.3 it is evident, that VOL is closely related to stand age. While up to stand age 60 a linear relationship can be observed, saturation starts at higher age.

A similar increase with stand age has also been observed for total biomass in boreal spruce \( (Picea abies) \) forests (Kazimirov & Morozova, 1973; cited from Schultz, 2000) but saturation did not occur before age 120. This relationship between stand age and total biomass was compared to the relationship between stand age and stem biomass (SBM). When SBM is plotted against stand age, apart from two extreme observations a curvilinear relationship is observed, but even at high age no saturation occurs (Figure 6.4). The outlier in the upper part of the plot is the oldest of all sampled stands (age = 148 years). In this stands very large values of canopy height and PBH and moderately large values of stem density were measured resulting in very large values of SBM. Despite a relatively large correlation coefficient between crown volume (VOL) and SBM of -0.74, VOL and SBM were considered as dissimilar variables in the subsequent analysis mainly due to their different dependence on stand age.
6.2.4 Relationships between vegetation indices and forest variables

Most of the linear regression models indicated that the best relationships were obtained using LAI or VOL; the regression with perimeter-at-breast height, canopy height, and stem density also provided several good relations but were not considered due to their strong intercorrelation; canopy cover was not further considered due to the strong relation to LAI. Thus, the results presented restrict to LAI, VOL, and SBM.

Narrowband VIs

To determine optimal narrowband VIs, coefficients of determination ($R^2$) between all possible two-band VIs and forest stand variables were computed. The results are illustrated in 2D-

![Figure 6.5: 2D-correlation plot that shows the correlation ($R^2$) between LAI and narrow band RVI values (subset $m$). The matrix is symmetrical, therefore just values above the diagonal are displayed. Below the diagonal, band combinations are marked in red where $R^2>0.75$. The displayed average reflectance spectrum of all measured forest plots eases the interpretation of the 2D-correlation plot.](image-url)
correlation plots (Figure 6.5). Each point at position (band 1, band 2) in the upper triangle of Figure 6.5 represents the $R^2$ value between LAI and the RVI calculated from the reflectance values in that two wavebands. The lower triangle highlights band combinations where $R^2$ was larger than 0.75. Similar correlation plots to the one in Figure 6.5 were computed for the other biophysical variables. Based on the $R^2$ values in the 2D-correlation plots band combinations that form the best indices were determined for LAI, VOL, and SBM. These were considered as optimal indices and were named RVI_opt and PVI_opt. Up to three best performing indices were considered when they occurred in different wavelength regions. For example, in Figure 6.5, the indices that have highest correlation with LAI were extracted from the 2D-correlation plot range above 0.75; the three most dominant indices occur in the regions around 1346nm/1207nm, 1134nm/920nm, and 1802nm/1163nm. The band positions were then tabulated in Table 6.7. For the best performing narrow band index, cross-validated $R^2$ and rmse were computed (Table 6.8).

**Broadband VIs**

The first step in the analysis of broadband VIs was to compare ratio to soil-adjusted VIs and

<table>
<thead>
<tr>
<th>Pooled dataset (t) (10-148 years)</th>
<th>Medium-old (mo) (30-148 years)</th>
<th>Old (o) (80-148 years)</th>
<th>Medium (m) (30-79 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$ [nm]</td>
<td>$R^2$</td>
<td>$\lambda$ [nm]</td>
<td>$R^2$</td>
</tr>
<tr>
<td>RVI_opt</td>
<td>918/965</td>
<td>.65</td>
<td>896/965</td>
</tr>
<tr>
<td>850/680</td>
<td>.55</td>
<td>1802/1043</td>
<td>.65</td>
</tr>
<tr>
<td>1165/750</td>
<td>.60</td>
<td>1346/1119</td>
<td>.65</td>
</tr>
<tr>
<td>PVI_opt</td>
<td>1088/1148</td>
<td>.70</td>
<td>895/1134</td>
</tr>
<tr>
<td>1148/807</td>
<td>.70</td>
<td>1320/1220</td>
<td>.82</td>
</tr>
<tr>
<td>RVI_opt</td>
<td>472/1119</td>
<td>.31</td>
<td>933/747</td>
</tr>
<tr>
<td>PVI_opt</td>
<td>777/900</td>
<td>.46</td>
<td>933/747</td>
</tr>
<tr>
<td>SBM</td>
<td>996/1058</td>
<td>.60</td>
<td>2385/2042</td>
</tr>
<tr>
<td>VOL</td>
<td>1417/1070</td>
<td>.82</td>
<td>1073/981</td>
</tr>
<tr>
<td>948/885</td>
<td>.82</td>
<td>2320/2449</td>
<td>.41</td>
</tr>
<tr>
<td>PVI_opt</td>
<td>671/702</td>
<td>.74</td>
<td>996/1058</td>
</tr>
<tr>
<td>981/1058</td>
<td>.82</td>
<td>1073/981</td>
<td>.60</td>
</tr>
</tbody>
</table>
to compare VIs based on visible and near-IR reflectance to those based on mid-IR and near-IR reflectance. Broadband ratio VIs (RVI and NDVI) generally showed relatively low values of $R^2$ and relatively high values of rmse for all subsets. The soil-adjusted broadband models (PVI and TSAVI) performed significantly better for the pooled data set; obviously, background effects related to soil and litter were reduced. Whereas younger stands revealed a denser canopy with little background contribution to the signal, in older stands with gaps and a more open canopy the background may have a larger influence on reflectance. Hence, when the total age dynamic is considered, TSAVI and PVI performed better than RVI or NDVI but when age classes are considered separately no performance increase was observed. The findings of Brown et al. (2000), that soil-adjusted VIs compared to the RVI have a decreased sensitivity to forest LAI, could not be confirmed.

Broadband MVI was closer related to LAI and VOL than broadband RVI and NDVI (subset m and mo). The mid-IR band in combination with the near-IR band seemed to contain more information relevant to the characterisation of forest canopies than the combination of red and near-IR bands. Recently, a closer relation of forest-LAI to radiation in the mid-IR than to radiation in the visible was found by Boyd et al. (2000) for tropical vegetation; the authors put forward that this could be the case also with boreal forests. The results found in the present study also support the suggestion by Fassnacht et al. (1997) that mid-IR bands may improve LAI estimation, particularly in more open forest stands. Recently, Lee et al. (2004) stretched the importance of spectral channels in the red-edge and mid-IR regions in predicting LAI of different biomes.

**Broadband versus hyperspectral VIs**

In the second step of the analysis, the broadband VIs were compared to the hyperspectral VIs to see if hyperspectral indices improve the prediction accuracy. All best narrowband RVI (RVI$_{opt}$) and PVI (PVI$_{opt}$) performed better than the corresponding broadband VIs. Values of rmse of about 1 m$^2$ m$^{-2}$ or larger for regression models between broadband RVI and LAI (medium-old ($mo$), old ($o$), and medium ($m$) aged stands) decreased to values as low as 0.5 m$^2$ m$^{-2}$ for the optimal narrowband RVI. The improvement of the narrowband models compared to the broadband models was not that distinct in the case of the PVI. Regression between PVI and VOL over all age classes (pooled dataset) revealed almost similar results for broadband and narrowband indices. The reason that some of the broadband results were high could be from having a very high signal-to-noise ratio (SNR). However, this does not mean that sensors like Thematic Mapper could be able to reproduce these results as they would have a much lower SNR.
Comparing narrowband orthogonal VIs with ratio VIs revealed that PVI_opt performed better
than RVI_opt (pooled dataset); for subsets \( m, o \), however, RVI_opt showed lower
rmse values than PVI_opt for LAI and VOL. No correlation was observed between the red
edge inflection point (REIP) computed from Lagrange interpolation (LGI) or inverted Gauss
model (IGM) and LAI over the total age dynamic, but for subset \( m \), LGI showed an rmse
value of less than 0.5 m\(^2\) m\(^{-2}\). This value is comparable to the rmse value obtained between
RVI_opt and LAI. These findings were partly supported by results obtained from other
studies. Danson & Plummer (1995) found a strong non-linear correlation between plot LAI
and the REIP for Sitka spruce (\( Picea sitchensis \)) using helicopter-borne spectroradiometer
data. For the same tree species, forest LAI was recently related to the canopy REIP computed
from imaging spectrometer data (CASI) with success (Lucas et al., 2000). Attempts to relate the chlorophyll content of canopies to the REIP were successful for grass canopies (Pinar & Curran, 1996), but only partially successful for forest canopies (Curran et al., 1991). Blackburn (2002) found no relation between REIP and LAI for coniferous stands using CASI data.

In a summary, it was concluded that the best relationships between remotely sensed reflectance and forest stand variables were found using LAI and crown volume (VOL). In a study on agricultural crops, Thenkabail et al. (2000) also found the closest relations between hyperspectral VIs and LAI or biomass, whereas relationships with crop height, and canopy cover were generally not as good. The major result of the present study was that the hyperspectral dataset contained more information relevant to the estimation of the forest stand variables LAI and VOL than multispectral data. Lee et al. (2004) came to the conclusion that regression models using AVIRIS channels performed better to predict LAI than those based on broadband data. However, the advantage of hyperspectral over multispectral data does not always seem to be the case: For instance, Broge & Mortensen (2002) came to the conclusion, that hyperspectral VIs derived from field spectral measurements were not better at estimating green crop area index (a variable related to LAI) than traditional broadband VIs. Then again, these authors found that the prediction of canopy chlorophyll density was improved using narrow bands across the red edge. Results of the present study concerning old stands show,
that relatively poor relationships were found in particularly for the broadband VIs. Also other studies reported problems with old stands that have been ascribed to shadow effects and a relatively dark background in the near-IR (Spanner et al., 1990a).

In the scatter plot between PVI_opt and LAI (Figure 6.6), even at high values of LAI no saturation is evident. A closer look reveals an outlier at position 4.1/-0.104. The corresponding forest stand of age 34 was identified in the image data. Whereas the 1999 image showed no abnormality in reflectance, in recently acquired HyMap data of 2003, a striped pattern was detected. From the spectral reflectance properties, the image pixels representing stripes could be identified as a mixture between tree crowns and forest litter. The striping pattern was caused by open strips that had been cut into the forest between the image acquisition and field measurements. Open strips allowed for the employment of harvesters in order to remove trees that had been exposed to game bite (Womelsdorf, G., personal communication).

Also the relationship between PVI_opt and VOL (Figure 6.7) is linear. In contrast to Figure 6.6, two main clusters of data points can be identified. Those data points with a value of VOL larger than 1 refer to forest stands of an age less than 40 years.
6.2.5 Absorption features related to leaf area index and crown volume

The best relations between both LAI and VOL and the optimal narrowband VIs used PVIs based on wavelength positions related to water absorption features. One of the two wavebands forming PVI_opt typically lay at the shoulder of the absorption feature, the other waveband at the absorption minimum (Figure 6.8). Obviously, PVI_opt based on such wavebands is a measure of the total amount of canopy water (M_{H_2O}) seen by the sensor. M_{H_2O} (unit: kg water per m^2 ground area) depends on both leaf water content (LWC; unit: mg water per cm^2 leaf area) and leaf area index (LAI):

\[
M_{H_2O} = LWC \times LAI
\]

When LWC does not vary between forest stands, PVI_opt reflects spatial patterns of LAI (Schlerf et al., 2005).

6.2.6 Maps of forest attributes

The best regression models that have been found for the estimation of LAI (Figure 6.9) and VOL (Figure 6.10) were applied to the HyMap image. Only image pixels representing conifer forest were considered. Estimated values of LAI were generally in a reasonable range. LAI values were classified into five classes for cartographic reasons. Although most of the forest stands in the LAI map appear rather homogenous, certain variation within the stands can be observed. This reflects the fact that even in highly managed forests stands do not develop evenly in space. From the interpretation of the spatial patterns prevalent in both parameter maps it can be seen that the values of the crown volume map are in general inversely related to the values of the LAI map. Young stands with low crown volume have not undergone clearing measures and reveal large LAI values. Old stands with relatively low values of LAI have built up large tree crowns. Medium old stands show the largest variability depending on to what extent clearing measures have been carried out.

From the resulting LAI map and the Forest Geographic Information System, average values
of LAI were computed for 156 forest stands (Figure 6.11). Stand average LAI was summarised into four age classes (Table 6.9). Mean LAI is generally decreasing with increasing stand age. All forest stands in the map support the age course of LAI observed for the 40 stands that have been used to calibrate the regression model (Figure 6.1) and that were then used to compute the map. Standard deviation of LAI is particularly large for the age class of 31-50 year old stands. This class also shows low minimum and large maximum values. The reason for the large variability of the 31-50 year old stands is that some stands of this age may be subjected to thinning measures while others may remain undisturbed and thinned at a later stage (Section 6.2.3).

The LAI map is a valuable resource. It was integrated into the local Forest-GIS of Idarwald. Stand LAI may serve to parameterise coarse scale ecological models (e.g., Biome-BGC) that are used to simulate net primary production and carbon, nutrient, and water cycling. The crown volume map may help to assess the amount of standing biomass in support of forest inventory procedures. Although remote sensing methods rarely play a role in operational inventory procedures, it may overcome certain limitations of traditional inventory practices.

6.2.7 Conclusions

This research intended to explore whether hyperspectral data may improve estimation of forest attributes compared to multispectral data using empirical statistical approaches. Several narrowband and broadband vegetation indices were compared. The following conclusions were drawn from this study:

- Forest leaf area index (LAI) and crown volume (VOL) were estimated with good accuracy from hyperspectral remote sensing data; stembiomass did not show relations with any vegetation index
- Orthogonal compared to ratio VIs were better suited to characterise forest LAI and crown volume
- Hyperspectral data contains more information relevant to the estimation of the forest stand attributes than multispectral data

Table 6.9: Summary statistics for leaf area index in five age classes

<table>
<thead>
<tr>
<th>Stand age [years]</th>
<th>n</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>11-30</td>
<td>24</td>
<td>5.2</td>
<td>0.7</td>
<td>3.3</td>
<td>6.8</td>
</tr>
<tr>
<td>31-50</td>
<td>46</td>
<td>3.7</td>
<td>0.9</td>
<td>1.9</td>
<td>6.7</td>
</tr>
<tr>
<td>51-80</td>
<td>41</td>
<td>3.1</td>
<td>0.6</td>
<td>2.2</td>
<td>4.6</td>
</tr>
<tr>
<td>≥ 81</td>
<td>45</td>
<td>2.6</td>
<td>0.4</td>
<td>1.7</td>
<td>3.9</td>
</tr>
</tbody>
</table>
Best hyperspectral VIs in relation with LAI were typically based on wavebands related to prominent water absorption features. However, more investigations are necessary to confirm this result.
Figure 6.9: Map of effective leaf area index at Idarwald obtained from an empirically derived predictive equation.
Figure 6.10: Map of crown volume at Idarwald obtained from an empirically derived predictive equation.
Figure 6.11: Map of stand mean leaf area index at Idarwald.
6.3 Inversion of a forest reflectance model

6.3.1 Introduction

The literature review (section 2.3.3) reveals that neural network approaches for the inversion of forest reflectance models using satellite or airborne remote sensing data that include a validation with ground truth data, are non-existing or rare. This investigation attempts to estimate forest structural characteristics from HyMap data through inversion of the hybrid-type reflectance model INFORM (Atzberger, 2000). The objectives were to i) assess if the reflectance signatures simulated by the model are consistent with the remotely sensed canopy spectra, ii) assess if the model can be inverted to provide reasonable estimates of forest leaf area index (LAI), crown cover and stem density, iii) identify the most suitable spectral bands for use in the neural network inversion procedure, and iv) and to compare inversion accuracies obtained with HyMap reflectances to those retrieved with HyMap data that had been resampled to Landsat TM wavebands.

6.3.2 Parameterisation of the forest reflectance model

A fundamental prerequisite to reflectance model inversion is a valid model that is capable of representing radiative transfer within the complex forest structure. To make sure that the model outputs were realistic, modelled reflectance was compared to reflectance measured by the HyMap sensor. Spectra were compared for canopy LAI (LAIc) classes of 1.5, 2.5, 3.5, 4.5, 5.5, and 6.5. The LAI map developed in section 6.2.6 was used to extract mean reflectance spectra from the HyMap imagery for a certain LAIc class value ± 0.5. For instance, for LAIc of 1.5, all pixels where the map showed LAI values from 1.0 to 2.0 were averaged to form a single reflectance spectrum. The LAI map had been previously derived through linear regression between a perpendicular vegetation index and field measured LAI at 40 plots; LAI had been estimated with a cross-validated $R^2$ of 0.77 (n = 39) and a cross-validated rmse of 0.54 m² m⁻² (17 percent of the mean) between estimated and measured LAI. The INFORM model requires the provision of a soil spectrum ($\rho_{soil}$). For that purpose a soil spectrum representing the forest floor (soil and litter, no understorey vegetation) was extracted from the HyMap dataset where the LAI map showed values around zero. The contribution of understorey vegetation is modelled by INFORM. A spectrum of infinite canopy depth was extracted from image pixels with LAI values above 10 according to the aforementioned LAI map. It well agreed with the modelled spectrum of infinite canopy depth.

In INFORM LAI is a single tree LAI (LAIₜ). Canopy LAI (LAIc) can be expressed as the product of LAIₜ and canopy cover (CO). CO was denoted “observed ground coverage by
crowns \( (c_o) \)" in section 4.4.2 (Equation 4.20) and is now simply referred to as “canopy cover”. CO is calculated as a function of stem density (SD) and crown diameter (CD) (Equations 4.20 and 4.22). To model forest reflectance of a certain LAI\(_c\), the values of LAI\(_s\), SD, and CD and also the canopy height (H) had to be defined. Site specific information was used to parameterise the model correctly and to model corresponding forest reflectance spectra. Field plot data (section 3.2.1) was used to establish relationships between LAI\(_c\) and attributes SD, CO and H (Table 6.10). Regression equations were used to compute values of SD, H and CO for the six LAI\(_c\) classes (Table 6.11). Single tree leaf area index (LAI\(_s\)) was computed as LAI\(_s\) = LAI\(_c\) / CO.

CD was measured only in forest stands older than 60 years \( (n = 28) \) and from this data no relationship could be found with LAI\(_c\). Typical values of CD ranged between 4 m for 60-80 year old trees up to 6.5 m for 80-100 year old trees. Generally, the size of tree crowns increases with increasing age and, due to the negative relationship between age and LAI\(_c\) found at Idarwald (section 6.2), it decreases with increasing LAI\(_c\). To represent this relationship and to extent it to forest stands younger than 60 years, values of CD ranging from 3.5 m to 6.5 m in steps of 0.6 m were assigned to the six LAI\(_c\) classes. All attribute values used to model canopy reflectance that were directly or indirectly derived from field work are listed in Table 6.11.

The remaining external, leaf and canopy parameters were kept constant for all simulations.

<table>
<thead>
<tr>
<th>LAI(_c)</th>
<th>LAI(_s)</th>
<th>SD [ha(^{-1})]</th>
<th>CO</th>
<th>CD [m]</th>
<th>H [m]</th>
<th>Modelled CO</th>
<th>Modelled LAI(_c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>4.4</td>
<td>261</td>
<td>0.34</td>
<td>6.5</td>
<td>35.2</td>
<td>0.41</td>
<td>1.78</td>
</tr>
<tr>
<td>2.5</td>
<td>5.9</td>
<td>395</td>
<td>0.42</td>
<td>5.9</td>
<td>29.5</td>
<td>0.48</td>
<td>2.81</td>
</tr>
<tr>
<td>3.5</td>
<td>7.0</td>
<td>597</td>
<td>0.50</td>
<td>5.3</td>
<td>23.8</td>
<td>0.55</td>
<td>3.82</td>
</tr>
<tr>
<td>4.5</td>
<td>7.8</td>
<td>902</td>
<td>0.58</td>
<td>4.7</td>
<td>18.0</td>
<td>0.61</td>
<td>4.75</td>
</tr>
<tr>
<td>5.5</td>
<td>8.4</td>
<td>1362</td>
<td>0.65</td>
<td>4.1</td>
<td>12.3</td>
<td>0.66</td>
<td>5.54</td>
</tr>
<tr>
<td>6.5</td>
<td>8.9</td>
<td>2058</td>
<td>0.73</td>
<td>3.5</td>
<td>6.5</td>
<td>0.70</td>
<td>6.19</td>
</tr>
</tbody>
</table>
Leaf variables as input to LIBERTY were set to default values as proposed by Dawson et al. (1998). Average leaf angle and variables related to the understorey were set to values as suggested by Atzberger (2000). The most important parameters are listed in Table 6.12.

Six forest reflectance spectra were simulated in 128 spectral HyMap wavebands using the INFORM model. As can be seen from Table 6.11, the modelled values of CO differed to a certain extent from those that were derived from field data through regression equations. If the field values of CO were taken as an absolute reference, the model overestimated small CO values and slightly underestimated large CO values. However, field values of CO were visually estimated and a relative, but no absolute accuracy could be expected. There was also a slight difference between modelled and field measured canopy LAI but it was generally smaller than for CO.

### Table 6.12: Important parameters that were kept constant during all model simulations. Type of variable: E = External, C = Canopy, L = Leaf

<table>
<thead>
<tr>
<th>Type of variable</th>
<th>Variable</th>
<th>Designation</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>Sun zenith angle</td>
<td>$\theta_s$</td>
<td>deg</td>
<td>30</td>
</tr>
<tr>
<td>E</td>
<td>Observation zenith angle</td>
<td>$\theta_o$</td>
<td>deg</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>Azimuth angle</td>
<td>$\psi$</td>
<td>deg</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>Fraction of diffuse radiation</td>
<td>skyl</td>
<td>fraction</td>
<td>0.1</td>
</tr>
<tr>
<td>C</td>
<td>Average leaf angle of tree canopy</td>
<td>ALA</td>
<td>deg</td>
<td>55</td>
</tr>
<tr>
<td>C</td>
<td>Leaf area index of understorey</td>
<td>LAI_U</td>
<td>m² m⁻²</td>
<td>0.1</td>
</tr>
<tr>
<td>C</td>
<td>Average leaf angle of understorey</td>
<td>ALA_U</td>
<td>deg</td>
<td>45</td>
</tr>
<tr>
<td>L</td>
<td>Cell diameter</td>
<td>d</td>
<td>µm</td>
<td>40</td>
</tr>
<tr>
<td>L</td>
<td>Intercellular air space</td>
<td>i</td>
<td>arbitrary</td>
<td>0.045</td>
</tr>
<tr>
<td>L</td>
<td>Leaf thickness</td>
<td>t</td>
<td>arbitrary</td>
<td>1.6</td>
</tr>
<tr>
<td>L</td>
<td>Baseline absorption</td>
<td>b</td>
<td>arbitrary</td>
<td>0.0006</td>
</tr>
<tr>
<td>L</td>
<td>Albino absorption</td>
<td>a</td>
<td>arbitrary</td>
<td>2</td>
</tr>
<tr>
<td>L</td>
<td>Chlorophyll content</td>
<td>c_{AB}</td>
<td>g m⁻²</td>
<td>200</td>
</tr>
<tr>
<td>L</td>
<td>Water content</td>
<td>c_W</td>
<td>g m⁻²</td>
<td>100</td>
</tr>
<tr>
<td>L</td>
<td>Lignin and cellulose content</td>
<td>c_L</td>
<td>g m⁻²</td>
<td>40</td>
</tr>
<tr>
<td>L</td>
<td>Nitrogen content</td>
<td>c_P</td>
<td>g m⁻²</td>
<td>1</td>
</tr>
</tbody>
</table>
CHAPTER 6 DETERMINATION OF FOREST STRUCTURAL ATTRIBUTES

Modelled and measured reflectance spectra are generally in good agreement (Figure 6.12). In the near-IR region both types of spectra show an increase in reflectance with increasing LAI\(_c\). Differences between modelled and measured spectra occur along the shoulder of the near-IR plateau, in the visible domain at large LAI\(_c\)-values and in the mid-IR wavebands at low values of canopy LAI. In the near-IR, both modelled and measured reflectances linearly increase with increasing LAI\(_c\) and no saturation effect can be observed (Figure 6.13, centre). Positive relationships between near-IR reflectance or NDVI and forest LAI or canopy cover were also

Figure 6.12: Measured (red) and modelled (blue) forest canopy reflectance for certain values of canopy LAI.

Figure 6.13: Measured (red) and modelled (blue) forest canopy reflectance against canopy LAI.
found in other studies that used forward reflectance modeling (Asner, 1998; McDonald et al., 1998; Eklundh et al., 2001). In the red, a decrease of reflectance with increasing LAI is expected from theoretical considerations which are confirmed by the modelled spectra (Figure 6.13, left). Here, saturation occurs at an LAI of about 4. Measured reflectance, in contrast, increases with increasing values of LAI in a linear fashion. The reason for this behaviour of the measured reflectances is that an increase of LAI comes along with a decrease of stand age and thus, less shadowing takes place. For the same reason, measured reflectance increases with increasing LAI in the mid-IR for values of LAI larger than 2 (Figure 6.13, right).

Up to an LAI of 4, modelled reflectance decreases with increasing LAI as expected but then increases for larger LAI values. Positive relationships between near-IR reflectance or NDVI and forest LAI or canopy cover were also found in other studies that used forward reflectance modeling (Asner, 1998; McDonald et al., 1998; Eklundh et al., 2001). For instance, model results using the Kuusk-Nilson model indicated that for dense forest stands, reflectance in the red decrease and in the near-IR increase with increasing LAI (Eklundh et al., 2001), but relations to measured reflectance were rather weak in all TM channels. However, in another study using the same model, negative relations between LAI and spectral reflectances in the red, near-IR and mid-IR were found (Rautiainen et al., 2004).

Figure 6.14: Modelled versus measured forest canopy reflectance for certain values of canopy LAI in different wavebands. The one-to-one line is shown for orientation.
Modelled were compared to measured reflectances in certain wavebands (Figure 6.14). The closest relation between modelled and measured reflectances occurred in the near-IR, followed by mid-IR and visible wavebands. Wavebands with best relationships were selected and used in the inversion process (section 6.3.7).

6.3.3 Overview of the inversion procedure

The general procedure for the estimation of forest attributes from remote sensing data through inversion of the reflectance model consisted of four major steps (Figure 6.15): Forward modeling to generate synthetic canopy spectra (a), training of the Artificial Neural Network (ANN) using the modelled data (b), application of the trained ANN to measured HyMap spectra to estimate forest canopy attributes (c), and the application of the ANN to the whole image to generate parameter maps (d). Details of the procedure are given in the following sections.

6.3.4 Forward reflectance modeling

Sets of input variables had to be generated to simulate forest reflectance spectra with INFORM. External parameters, leaf parameters and certain canopy parameters were assigned fixed standard values as defined in Table 6.12. Values of the canopy variables LAI₅₅, SD, CD,
and $H$ were chosen randomly within the range of values observed at the study site. Their statistics are given in Table 6.13. 5000 sets of input variables were generated and the corresponding forest reflectance spectra were modelled in 128 HyMap wavebands.

### 6.3.5 Spectral sampling

Given a multi- or hyperspectral remote sensing data set, a selection of the bands to be used in the inversion process is recommended because those bands with larger model and measurement uncertainties would bias the retrieval (Meroni et al., 2004). However, an optimal spectral sampling is a very complex problem and is still an open issue for the remote sensing community. Recently, Genetic Algorithms have been used to find the optimal set of spectral variables to include into a model.

To compare the suitability of broadband multispectral and narrowband hyperspectral data in model inversion, simulated and measured HyMap reflectances were spectrally resampled to Landsat TM wavebands using the appropriate sensor functions. We tested all two-band and three-band combinations in TM wavebands. More than three wavebands were not considered to avoid usage of highly correlated data. It turned out that two wavebands were sufficient to invert the model. To save computation time, in HyMap wavebands just all two-band combinations were assessed, which meant to 114x113/2 times invert the model.

### 6.3.6 Network architecture and training

The Neural Network approach was realised within the Matlab software environment (www.mathworks.com). Training of the network was done by presenting modelled spectra and canopy variables to the ANN. During the training process, the values of the weights and biases in the network are iteratively adjusted in such a way that the error between the network outputs and the target outputs (the canopy variables) is minimised. Before the modelled canopy variables and spectra could be presented as training data to the ANN, two preprocessing steps were performed: Band selection and data normalisation.
To ensure that the training data was consistent with the real data, wavebands were selected where measured and modelled reflectances agreed best (Figure 6.14). Selection of input bands involved testing all 2-band combinations of wavebands at 899.6 nm, 1073.5 nm, 1278.7 nm, and 1669.4 nm on how well the network performed on the validation data (section 6.3.7). From the available wavebands, two were selected: one in the near-IR at 1073.5 nm and another one in the mid-IR at 1669.4 nm. More than two wavebands were not considered to avoid usage of highly correlated data. Spectra and canopy variables were normalised by a forward z-transformation. Statistical parameters of the z-transformation (mean and standard deviation) were later used to normalise the measured data in exactly the same way.

The 5000 sets of reflectance data in the selected wavebands together with the corresponding input variables were divided into a training dataset and a test dataset. The measured HyMap spectra and the field-measured canopy variables were used as validation dataset. All following assessments concerning the performance of the ANN are based on the results obtained with the validation set. Details of the validation procedure are given in section 6.3.7. The structure of the used ANN was relatively simple (Figure 6.16). It had two input neurons (equal to the number of used wavebands), one neuron in the hidden layer, and three output neurons (one for...
each variable to be estimated). A three layer feed-forward backpropagation network with a
tan-sigmoidal transfer function in the hidden layer and a linear transfer function in the output
layer was employed. The layer with the nonlinear transfer function allowed the network to
learn nonlinear and linear relationships between input and output vectors. The linear output
layer enabled the network to generate values outside the range -1 to 1. The employed network
type can approximate any function, given sufficient neurons in the hidden layer (Demuth &
Beale, 2003).

For training the resilient backpropagation algorithm was applied. The purpose of this training
algorithm is to eliminate harmful effects of the magnitudes of the partial derivatives that may
be caused when using steepest descent to train a multilayer network with sigmoid functions.
Only the sign of the derivative is used to determine the direction of the weight changes; the
magnitude of the derivative has no effect on the weight changes (Demuth & Beale, 2003).
Other faster training algorithms, the Levenberg-Marquardt backpropagation and the scaled
conjugate gradient backpropagation algorithm were also tested but not further considered, as
the resilient backpropagation algorithm produced the best results on the validation data.

One of the main problems that typically occur during ANN training is overfitting. This
phenomenon describes the situation that the error on the training data is driven to a very small
value, but when new data is presented to the network the error is large (Demuth & Beale,
2003). This lack of generalisation can be avoided by using a network that is just large enough
to provide an adequate fit. The complexity of a network is mainly determined by the number
of neurons in the hidden layer. To find the appropriate dimension of the hidden layer, the
performance of the algorithm was systematically investigated with the number of neurons in
the hidden layer varying between 1 and 10; it turned out that one neuron in the hidden layer
was sufficient to approximate the function between input and output vectors of the validation
data.

To further improve generalisation of the network the early stopping technique (Demuth &
Beale, 2003) was applied. The patterns generated with INFORM were divided into two
subsets. 2500 datasets were used for training, and 2500 were used as a test dataset to prevent

Figure 6.17: Modelled against measured forest canopy variables (n = 39).
overfitting. The training set was used for computing the gradient and updating the weights and biases. The error on the test set was monitored during the training process. During the initial training phase, both, training and test set error decreased. When the network began to overfit the data, the test set error started to rise. At this point the training was stopped and the actual weights at the minimum of the test error were returned. These weights and biases were then used to estimate the independent variables of the validation dataset.

### 6.3.7 Validation of inversion results and modification of network architecture

Only data from 39 plots were used for validation since one of the originally 40 plots had been thinned between the image acquisition and field measurements (section 6.2.4). The validation data (i.e. the HyMap spectra and the structural attribute data) was preprocessed in exactly the same way as the training set. That means the same wavebands were selected from the measured HyMap spectra. Both reflectances and attribute data were normalised with the same statistical parameters that had been previously used for the z-transformation of the training data. After the HyMap spectra were sent through the network, the resulting network outputs were subjected to an inverse z-transformation and from the estimated values of LAI, SD, and CD, estimated values of CO and LAI$_c$ were computed using equations 4.20 and 4.22 (Figure 6.16). The coefficient of determination and the root mean squared error between estimated and field-measured forest canopy variables were calculated. Both parameters served as a measure of accuracy of the model inversion. The number of hidden neurons and input bands were systematically analysed (section 6.3.6). To generate biophysical parameter maps the trained ANN yielding the highest overall accuracy was finally applied to the radiometrically and geometrically corrected HyMap image.

Using two HyMap wavebands at 837 nm and 1148 nm, model estimated and ground LAI show relatively good agreement as indicated by the close position of most data points along 1:1 line (Figure 6.17). Low values of ground LAI are slightly overestimated whereas relatively large values of ground LAI are somewhat underestimated. The obtained accuracy of the LAI map amounts to an rmse of 0.58 (relative rmse 18 % of mean) and a coefficient of determination ($R^2$) of 0.73. As expected, canopy cover and stem density (both considerably determine canopy LAI in the INFORM model) behave similar with LAI with respect to the 1:1 line. Canopy cover was estimated with reasonable accuracy (rmse 17.5 % of mean). However, model outputs of stem density considerably underestimate the ground measurements for large values (rmse 55 % of mean). With HyMap data resampled to Landsat TM spectral bands and using two bands at 840 nm and 1650 nm, rmse was 0.66 and relative rmse 21 % against ground measured LAI.
Figure 6.18: Map of effective leaf area index at Idarwald obtained through inversion of a forest reflectance model.
Figure 6.19: Map of stem density at Idarwald obtained through inversion of a forest reflectance model.
The main result of this research is that forest leaf area index could be estimated through inversion of a reflectance model with accuracy comparable to that of VI based approaches. These findings are supported by other studies. For instance, LAI of different vegetation types was estimated with fair accuracy (rmse = 0.70) by inversion of the Kuusk reflectance model, better than that obtained with an empirical LAI-VI relation (Bicheron & Leroy, 1999). In another work, inversion of the hybrid-type Kuusk-Nilson model performed better than NDVI when estimating forest cover from TM data, though accurate estimates on a per-stand basis were not possible (Gemmell et al., 2002). Whereas empirical approaches rely on a calibration function between a certain vegetation index and a ground measured forest variable, calibration of the reflectance model is based on synthetic data and the ground truth serves as an independent validation base. In opposition to empirically derived prediction functions that are generally limited to the local conditions at a certain point in time and to a specified sensor type, the calibrated reflectance model can be applied more easily to different optical remote sensing data acquired at landscape to regional scales (Schlerf & Atzberger, 2006).

6.3.8 Maps of forest attributes

The final neural network was applied to the entire HyMap image. Again, only image pixels representing conifer forest were considered. Estimated values of LAI were generally in a reasonable range (Figure 6.18). LAI values were classified into five classes for cartographic reasons. Stem density is presented in values ranging from low to high as estimated absolute values did not well agree with those that had been measured in the field (Figure 6.19).

6.3.9 Conclusions

If correctly parameterized, the INFORM model simulates forest reflectance spectra comparable to those measured by an imaging spectrometer (HyMap). There is a good agreement between modelled and measured spectral reflectance in the near-IR wavebands and an acceptable agreement in the mid-IR. In the visible, however, model outputs and HyMap reflectances disagree and the model has to be further improved. Our results further show that INFORM can be inverted with a neural network approach to give estimates of important structural forest characteristics. Relatively simple network architecture with just one hidden neuron proofed to be suitable to solve the inversion problem. By systematically evaluating all 2-band combinations of HyMap spectral bands, the optimum band combination could be identified. Estimates of forest LAI and percent coverage show relative good agreement with ground truth data (n = 39). In contrast, the spectrally much less important stem density could only be estimated with fair quality. Using HyMap reflectances resampled to Landsat TM spectral bands it could be shown, that two spectral bands were sufficient to properly invert the model; usage of more spectral bands did not further improve the inversion. Compared to the results using the HyMap data, Landsat TM performed less well.
Chapter 7  Estimation of forest chemical attributes

7.1 Introduction

This chapter focuses on the estimation of foliar concentrations of chlorophyll, nitrogen and water from high spectral resolution data. The research systematically compared predictive models between different types of spectral data and foliar chemical concentrations. It is based on the spectral reflectance measurements (section 3.2.3) and chemical analyses (section 3.2.2) of Norway spruce needles collected at Gerolstein and Idarwald test sites (section 3.2.1) and the HyMap image acquired over the Gerolstein area in 2003 (section 3.3.2). Whereas chemical data from Gerolstein was used to develop predictive models, chemical data from Idarwald was used to test the predictive models. In a first step, the approach was tested on needle spectra (layer stack) measured in the laboratory and afterwards applied to forest canopy spectra measured by the HyMap sensor.

The overall aim of the work was to evaluate the information content of hyperspectral remote sensing data for the estimation of foliar chemical concentrations in forests. More specific objectives were (i) to find predictive models by stepwise multiple linear regression technique (section 4.3.2) using characteristic absorption features to estimate chemical concentrations, (ii) to systematically compare different types of derivatives of spectral reflectance regarding the accuracy of the predictive models and (iii) to examine the effects of sensor bandwidth and sensor noise on the estimations of chemical concentrations from laboratory reflectance data to assess the transferability of the techniques to remote sensing data.

Different types of derivatives of spectral reflectance have been used in combination with stepwise multiple regression (SMR) to estimate foliar chemical concentrations, such as first-derivative of reflectance (section 4.3.6), continuum-removed (CR) spectra and band-depth normalised (BDN) spectra (section 4.3.7) among others. Most of these derivatives aim to enhance the absorption feature present in vegetation spectra in the first step and SMR in the second step tries to find wavebands located in these absorption features that are related to the investigated chemicals. CR and BDN spectra were most often applied to reflectance spectra of dry leaf samples or to absorption features that are not masked by water (section 2.3.4), for instance the chlorophyll features. A technique that explicitly attempts to remove the influence
of leaf water from a fresh vegetation reflectance spectrum is least-squares spectral matching (LSM). LSM was modified (section 4.3.8) and combined with SMR to investigate if it allows for a better estimation of leaf nitrogen than the other derivatives mentioned before. The following types of derivative of reflectance were compared:

- Spectral reflectance (RFL)
- First-difference of spectral reflectance (DRFL)
- Continuum-removed spectra (CR)
- Band-depth normalised spectra (BDN)
- First-difference of water removed spectra (DWRS)

The following leaf chemical concentrations were investigated:

- Chlorophyll a+b concentration ($C_{AB}$)
- Nitrogen concentration ($C_N$)
- Water concentration ($C_W$)

The next section briefly explains the general procedure of developing predictive models, that is, how the absorption features relevant to the determination of the chemicals were extracted and how the SMR method and the validation procedure were implemented. The third section shows the results obtained with the modified LSM technique to derive an additional derivative of reflectance and how it was coupled with SMR. The fourth section shows the results obtained with the laboratory spectra and the fifth section shows the results achieved with the HyMap image. The chapter ends with the conclusion drawn from this study.
Summary statistics of chemical data from both test sites are listed in Table 7.1. From the summary statistics, it can be seen that $C_{AB}$ has a larger variability than $C_N$ and $C_W$ at both test sites. Mean values of $C_{AB}$ and $C_W$ were significantly larger at Gerolstein than at Idarwald whereas mean $C_N$ showed no marked difference. First year needles have higher water and concentrations but lower chlorophyll $a+b$ concentrations then third year needles. Nitrogen concentration does not show a marked difference between first and third year needles.

A complete list of the chemical analysis data is given in Appendix B. Linear correlation coefficients between $C_{AB}$ and $C_W$ ($n=78$) indicate a low negative relation ($r=-0.39$) and between $C_{AB}$ and $C_N$ a moderate positive relation ($r=0.58$). The closest relation is found between $C_N$ and $C_W$ ($r=0.73$). To examine the potential of the method at the fresh leaf level to cope with remote sensing measurements of forest canopies, the influence of sensor bandwidth and sensor noise were considered. To do so, two additional sets of reflectance were computed from the reflectance measured by the Field-Spec-II device (F-data). Reflectance spectra of 1 nm spectral resolution were degraded to HyMap spectral bands using the appropriate sensor function (FH-data). To the degraded spectra a normally distributed noise component ($\sigma=0.002$) was added (FHN-data).

![Figure 7.1: Location of absorption features related to total chlorophyll (a), water (b), and protein (c, d, e) used in the regression analysis.](image-url)
7.2 Developing predictive models

7.2.1 Selection of absorption features

Kokaly & Clark (1999) used three absorption features located around 1730 nm, 2100 nm, and 2300 nm to estimate nitrogen, lignin, and cellulose concentrations. Curran et al. (2001) increased the number of chemicals under investigation and included additional absorption features into their analysis. Following both approaches, the chlorophyll absorption feature at 670 nm and the water absorption feature at 1200 nm, and three absorption features related to protein at 1800 nm, 2100 nm, and 2300 nm were included into the regression analysis in order to estimate concentrations of nitrogen, chlorophyll and water (Table 7.2, Figure 7.1). Calculation of the derivatives involved either computation of the derivative followed by an extraction of the features (RFL, DRFL) or extraction of the features followed by the computation of the derivative (CR, BDN). The five features of a single type of derivative were then re-combined into one single data set, normalised band wise to mean of zero and standard deviation of one, and subjected to regression analysis. SMR was run separately for 1st year, 3rd year, and both 1st and 3rd year needle samples. The idea was that each data set contained all information necessary to estimate all of the three chemical substances and to see if the appropriate wavebands were selected from the available bands.

7.2.2 Developing regression equations

Forward SMR was used to develop relationships between chemical concentration and a

![Figure 7.2: Cross-validated rmse against number of wavebands selected by stepwise multiple regression for the Field-Spec-II data resampled to HyMap spectral bands (DWRS data set). Some number of bands did not occur as a consequence of the used enter and removal criteria. DWRS refers to first-difference of water removed spectra (see section 7.3).](image)
measure of reflectance for certain wavebands. Cross-validation was employed to assess the accuracy and validity of the regression models. This required for each regression variant to develop n (n = sample number) separate models, each time with data from n-1 observations. The calibration model was then used to predict the observation that was left out. Cross-validated coefficient of determination ($R^2$) and cross-validated root mean squared error (rmse) as well as relative rmse (rmse divided by the mean) provide performance measures for each chemical. Forward SMR was run repetitively with multiple pairs of $p$-values to enter ($p_{in}$) and remove ($p_{out}$) on the spectral data and the chemical data. Larger $p$-values often resulted in a larger number of wavebands included into the regression models. With increasing number of wavebands in the regression model, overfitting of the model can occur (Curran, 1989). If overfitting occurs, SMR may find false patterns in the data. For instance, Grossman et al. (1996) found SMR relating chemistry data and derivative reflectance spectra to give relatively large coefficients of determination even for randomised data sets. To avoid overfitting of the models, two control measures were implemented: i) cross validated values of rmse were plotted against the number of bands selected (Figure 7.2). Optimal numbers of bands were determined visually by only including a maximum number of bands where the rmse values were above the break in slope. For instance, for the estimation of CW from DWRS data (year 1 needles), three wavebands were considered to be appropriate. An increase in the number of selected bands would not have noticeably decreased rmse. Entry and removal criteria were modified in such a way that the optimal number of bands that had been previously determined was included into the model. ii) Control runs were done with chemical data that had been brought into random order. When a control run for a particular model achieved values of $R^2$ larger than zero, $p$-values for entry and removal were decreased in such a way, that no correlation between spectral data and randomized chemical data was observed any longer. These criteria were then used in the actual model run. After taking into account both control measures, models typically included between 2 and 7 wavebands in case of the needle spectra (n=78 for 1st and 3rd year samples) and between 1 and 3 wavebands in case of the canopy spectra (n=13).
7.3 Removing water from absorption features in the mid infrared

7.3.1 Testing the modified least squares spectral matching approach

In reliably estimating leaf nitrogen concentration from fresh vegetation spectra, as main limiting factors have to be considered the leaf water that obscures absorption features related to protein and other chemical substances. If a refined methodology was developed to remove the disturbing effects of leaf water on the estimation of protein from reflectance data then it would have potential to better estimate leaf nitrogen concentration. Such a technique, least-squares spectral matching (LSM), was published by Gao & Goetz (1994, 1995, section 4.3.8).

Figure 7.3: A) Matching of fresh needle (stack layer) spectrum in the 1500-1780 nm region with the water spectrum and with a combination of the water and the protein spectrum, B) residual spectra between the measured and the modeled spectra.
To test the LSM technique on the measured laboratory spectra, the matching was made over a narrow spectral range between 1.50 µm and 1.78 µm located between the major atmospheric water absorption bands as suggested by Gao & Goetz (1994). To compare the goodness of the matching, the residual spectrum (Residual of Spectral Matching, RSM) was calculated which is defined as the difference between the measured fresh needle spectrum $R_{F_{mes}}$ and the modelled spectrum $R_{F_{mod}}$ or $R_{W_{mod}}$ divided by the measured spectrum $R_{F_{mes}}$ (Figure 7.3). The dashed line shows the matching of the measured spectrum of fresh spruce needles with the leaf water spectrum $R_{W_{mod}}$. Obvious differences between both spectra occur near 1570 nm, 1650 nm, and 1725 nm. The dotted line in Figure 7.3 shows the match of the measured fresh needle spectrum $R_{F_{mes}}$ with a combination of the leaf water and the protein spectrum $R_{F_{mod}}$. The values of RSM at 1650 nm and 1725 nm are considerably decreased for the fresh needle spectrum compared to the water spectrum. Obviously, the reflectance spectra of spruce needles or green vegetation in general in the 1.50 µm to 1.78 µm region are both controlled by leaf water and protein. The remaining differences between the modelled fresh needle spectrum and the measured spectrum can probably be attributed to other chemicals, such as lignin and cellulose, as they also influence the reflectance spectra in that spectral region (Gao & Goetz, 1994). It could be followed that the residual of the matching of the fresh needle spectrum with the leaf water spectrum in the 1.50-1.78 µm spectral region contains information about the needle’s protein concentration. As many protein absorption features are located in the 2.1-2.3 µm spectral region it was assumed that the residual in the 2.1-2.3 µm spectral region could also contain relevant information about protein. In a next step the RSM spectra were subjected to a stepwise regression procedure.

<table>
<thead>
<tr>
<th>Spectral data</th>
<th>Leaf chemical</th>
<th>$R^2$</th>
<th>rmse</th>
<th>Bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRFL</td>
<td>$C_N$</td>
<td>0.29</td>
<td>0.10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>$C_W$</td>
<td>0.53</td>
<td>7.37</td>
<td>3</td>
</tr>
<tr>
<td>DWRS</td>
<td>$C_N$</td>
<td>0.54</td>
<td>0.08</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>$C_W$</td>
<td>0.06</td>
<td>10.49</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 7.3: The cross-validated coefficient of determination ($R^2$) and the cross-validated rmse between two sets of spectral data (features at 1800 nm, 2100 nm, and 2300 nm) and nitrogen and water for 1st year needle reflectance spectra resampled to HyMap spectral bands (n=39); bands: number of wavebands selected by SMR ($p_{in} = 0.10$, $p_{out} = 0.15$); units of rmse: % dry matter.
7.3.2 Coupling least squares spectral matching with stepwise multiple regression

LSM was run on the investigated absorption features at 1800 nm, 2100 nm, and 2300 nm using equation 4.16 to produce water removed spectra (WRS). To examine if LSM could remove the influence of water on reflectance spectra in this wavebands, SMR was run on the first difference of WRS (DWRS) and the first difference of reflectance (DRFL) data and concentrations of nitrogen and water. First difference of reflectance approximates the slope of reflectance (first derivative) which has a closer relation to absorption features than absolute values of reflectance (Dixit & Ram, 1985). Results of the analysis (Table 7.3) showed that for 1st year needles the $R^2$ between DRFL and water was larger than between DRFL and nitrogen. However, after the application of LSM to the spectra, the $R^2$ between DWRS and nitrogen was larger than between DWRS and water. The $R^2$ between spectral data and nitrogen increased from 0.29 with DRFL data to 0.54 with DWRS data. For 3rd year needles the difference between DRFL and DWRS in estimating nitrogen was less pronounced. Nevertheless, it was concluded that LSM can remove the influence of leaf water on the reflectance in relevant wavebands of the mid-IR region. At the same time, LSM may enhance absorption features related to nitrogen containing leaf constituents such as protein.

Table 7.4: The cross-validated coefficient of determination ($R^2$) and the cross-validated rmse between spectral data (DWRS) and concentrations of chlorophyll, nitrogen and water (year 1: n=39, year 3: n=39, year 1+3: n=78); bands: number of wavebands selected by SMR; units of rmse: % dry matter ($C_W$, $C_N$), mg g$^{-1}$ dm ($C_{AB}$)

<table>
<thead>
<tr>
<th></th>
<th>Field-Spec-II data (F-data)</th>
<th>Field-Spec-II data resampled to HyMap spectral bands (FH-data)</th>
<th>Field-Spec-II data resampled to HyMap spectral bands plus noise (FHN-data)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>707 wavebands</td>
<td>49 wavebands</td>
<td>49 wavebands</td>
</tr>
<tr>
<td></td>
<td>$R^2$          rmse   Bands</td>
<td>$R^2$          rmse   Bands</td>
<td>$R^2$          rmse   Bands</td>
</tr>
<tr>
<td>Year 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_{AB}$</td>
<td>0.82            0.19   3</td>
<td>0.81            0.20   3</td>
<td>0.69            0.24   3</td>
</tr>
<tr>
<td>$C_W$</td>
<td>0.74            5.5    6</td>
<td>0.60            6.8    3</td>
<td>0.32            8.9    3</td>
</tr>
<tr>
<td>$C_N$</td>
<td>0.76            0.059  6</td>
<td>0.63            0.074  6</td>
<td>0.44            0.089  3</td>
</tr>
<tr>
<td>Year 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_{AB}$</td>
<td>0.68            0.32   3</td>
<td>0.63            0.34   3</td>
<td>0.66            0.33   3</td>
</tr>
<tr>
<td>$C_W$</td>
<td>0.73            5.2    7</td>
<td>0.63            6.1    7</td>
<td>0.26            8.56   2</td>
</tr>
<tr>
<td>$C_N$</td>
<td>0.82            0.043  7</td>
<td>0.72            0.053  5</td>
<td>0.54            0.068  2</td>
</tr>
<tr>
<td>Year 1+3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_{AB}$</td>
<td>0.83            0.26   3</td>
<td>0.81            0.28   3</td>
<td>0.75            0.32   3</td>
</tr>
<tr>
<td>$C_W$</td>
<td>0.78            8.0    3</td>
<td>0.77            8.1    3</td>
<td>0.57            11.1   2</td>
</tr>
<tr>
<td>$C_N$</td>
<td>0.56            0.076  4</td>
<td>0.52            0.080  4</td>
<td>0.47            0.083  2</td>
</tr>
</tbody>
</table>
To compare DWRS with the other derivatives spectral data representing the chlorophyll and the water absorption features had to be used. However, for these two features water removal did not make sense (as no water effect is present in the chlorophyll absorption feature at 670 nm and water influence was not to be removed from the main water feature at 1200 nm). Thus, for the chlorophyll and water absorption features, first derivative of reflectance (DRFL) and for the protein absorption features, first derivative of water removed spectra (DWRS) were used and combined into a single data set. The dataset was denoted DWRS. By this, all spectral datasets consisted of the same number of bands (F-data: 707 wavebands, FH-data and FHN-data: 49 wavebands).

Figure 7.4: Positions of wavebands that have been selected by SMR between DWRS spectral data (resampled to HyMap wavebands) and three leaf chemicals.
7.4 Estimating leaf chemicals from laboratory reflectance data

7.4.1 Effects of sensor bandwidth and sensor noise

The first part of the analysis exemplarily used DWRS spectra to compare the performance of laboratory data (F-data), laboratory data resampled to HyMap spectral bands (FH-data), and noisy resample laboratory data (FHN-data) (see introduction in section 7.1) in estimating foliar chemicals. Cross-validated coefficients of determination ($R^2$) for all samples (Year 1+3) were generally largest between F-data and total chlorophyll (C$_{AB}$), second largest between F-data and water (C$_{W}$) and lowest for F-data and nitrogen (C$_{N}$) (Table 7.4). $R^2$ were generally largest between F-data and chemical concentrations and lowest for FHN-data and chemical attributes. Inversely, cross-validated values of rmse were generally lower between F-data and

Table 7.5: Wavebands that have been selected by SMR between DWRS spectral data (resampled to HyMap wavebands) and three leaf chemicals; bold: wavebands directly related to known absorption features of the chemical of interest; italic: wavebands related to known absorption features of a chemical with which the chemical of interest was correlated (indirectly related); Interpretation consists of the related absorption feature (if present) and its location (in nm) or describes the spectral region where the selected waveband is located; Wavebands appear in the order of selection by SMR

<table>
<thead>
<tr>
<th>Leaf chemical</th>
<th>Year</th>
<th>Selected wavebands [nm] and interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorophyll a+b</td>
<td>1</td>
<td>648.5: Chlorophyll b at 640, 739: Red edge, 1126.5: Lignin at 1120</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>723.5: Red edge, 739: Red edge, 1141: Lignin at 1120</td>
</tr>
<tr>
<td></td>
<td>1+3</td>
<td>601.5: Edge of chlorophyll absorption feature, 739: Red edge, 1155.5: Edge of water absorption band</td>
</tr>
<tr>
<td>Chlorophyll a+b</td>
<td>1</td>
<td>1126.5: Lignin at 1120, 1229: Water at 1200, 1737: Protein at 1730</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>571.5: Edge of chlorophyll absorption feature, 601.5: Edge of chlorophyll absorption feature, 648.5: Chlorophyll b at 640, 663.5: Chlorophyll a at 660</td>
</tr>
<tr>
<td></td>
<td>1+3</td>
<td>1155.5: Edge of water absorption band, 1229: Water at 1200, 2210.5: Protein at 2240</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>1</td>
<td>556: Green peak, 648.5: Chlorophyll b at 640, 739 nm: Red edge, 1700: Protein at 1690, 1725: Protein at 1730, 2362: Protein at 2350</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>632.5: Chlorophyll b at 640, 739: Red edge, 1126.5: Lignin at 1120, 2192.5: Protein at 2180, 2262.5: Protein at 2240</td>
</tr>
<tr>
<td></td>
<td>1+3</td>
<td>648.5: Chlorophyll b at 640, 739: Red edge, 1214.5: Water, Cellulose, Starch, Lignin at 1200, 2262.5: Protein at 2240</td>
</tr>
</tbody>
</table>
foliar chemicals and largest for FHN-data and leaf constituents. However, the effects of spectral resolution and noise were not the same for all chemical attributes. A decrease of spectral resolution had a minor influence on the estimation of both $C_{AB}$ and $C_W$. Adding noise to the data, on the other side, strongly decreased the accuracy for the estimation of $C_W$, but only slightly decreased the accuracy for the estimation of $C_{AB}$.

Estimation accuracy for nitrogen ($C_N$) decreased from F-data over FH-data to FHN-data for Year 1 and Year 3 needles but remained almost constant for Year 1+3 samples. The number of bands included into the models was only slightly lower for FH-data than for F-data, but considerably lower for FHN-data. Whereas models to estimate $C_{AB}$ always included 3 wavebands, the number of bands included into a model considerably decreased after adding noise to the data in the case of $C_W$ and $C_N$. In a summary, when looking on all samples ($n=78$), $C_{AB}$ was well estimated and just slightly affected by a decrease of the spectral resolution or adding of noise, $C_N$ was moderately well estimated and just slightly affected by a decrease of the spectral resolution or adding of noise and $C_W$ was well estimated with F-data and FH-data but strongly affected by noise addition. For subsequent analyses, only FH-data were considered.

### 7.4.2 Selected wavebands by stepwise multiple regression

The wavebands selected by stepwise multiple regression (SMR) between FH-data (DWRs spectra) and concentration of foliar chemicals are presented in Table 7.5 and Figure 7.4. Of the wavebands selected to estimate $C_{AB}$, only one was directly related, only two were indirectly related, and six were not related to known absorption features listed in Table 2.1. Wavebands located in the red edge region and along the edge of the chlorophyll absorption feature played an important role in the estimation of $C_{AB}$ concentration from spectral data. This result is in agreement with a general understanding that the maximum correlations occur on the edge of absorption features as the central wavelengths become saturated (O’Neill et al., 2002; Curran, 1989). Of the wavebands selected to estimate $C_W$, three were directly related, six were indirectly related, and four were not related to known absorption features (Table 7.5). Regression models to estimate $C_W$ often employed wavebands related to $C_{AB}$ (year 3) or other chemicals (year 1) or made use of wavebands located at the edge of water absorption features (year 1+3). Of the wavebands selected to estimate $C_N$, nine were directly, two indirectly, and four not related to known absorption features. The first waveband selected in a regression model was often related to the chlorophyll absorption feature and the second was always located in the red edge region, reflecting the fact that chlorophyll is a major nitrogen bearing leaf constituent. The wavebands selected afterwards were often directly related to known protein absorption features as protein is the second most abundant constituent containing nitrogen.
The results demonstrate that nitrogen concentration can be moderately well estimated from reflectance spectra using wavebands related to chlorophyll and protein absorption features. Previous studies have also used statistical methods to link reflectance in certain wavebands to changes in nitrogen, protein or chlorophyll concentration. Table 7.6 presents the selection of wavebands in previous studies. Wavelengths close to protein features at 1730 nm, 2180 nm, and 2350 nm that have been repeatedly selected in these studies were also selected in the present work (Table 7.5). As in the present work, a combined selection of both, wavebands related to chlorophyll and protein absorption features in one model to estimate nitrogen concentration, had also occured in previous studies (Johnson et al., 1994; Curran et al., 2001; Serrano et al., 2002).

Table 7.6: Wavelengths correlated with nitrogen (Nit), protein (Pro), total chlorophyll (Chl) and water (Wat) by multivariate statistical measures; Level: D = Dried, ground leaves, F = Fresh, whole leaves, CA = Canopy measured from airborne sensor; Spectral Variable: RFL = Reflectance, DRFL = First derivative of reflectance, NBA = Normalised to band area; bold: wavebands directly related to known absorption features of the chemical of interest; italic: wavebands related to known absorption features of a chemical with which the chemical of interest was correlated (indirectly related)

<table>
<thead>
<tr>
<th>Authors</th>
<th>Species</th>
<th>Level</th>
<th>Spectral Variable</th>
<th>Wavelengths [nm]</th>
</tr>
</thead>
</table>
| Card et al., 1988   | Various tree species     | D     | Log (1/RFL)       | Pro: **2220**, 2110, 1730, 1460, 1740, 2170
|                     |                          |       |                   | Chl: **680**, 2110 |
| Wessman et al., 1989| Deciduous                | CA    | DRFL              | Nit: 1265, 1555  |
| Curran et al., 1992a| Amaranthus               | F     | DRFL              | Pro: **2322**    |
|                     |                          |       |                   | Chl:: 724       |
|                     |                          |       |                   | Wat: 1752, 1268, 550, 2148, **1172** |
| Johnson et al., 1994| Various conifer species  | CA    | DRFL              | Nit: 552, **1184**, **2164** |
|                     |                          |       |                   | Chl: 1069, **1683**, 858 |
| Curran et al., 1997 | Slash pine               | CA    | DRFL              | Nit: **1031**, **925**, **1223**, **2341**, **916** |
| Kokaly & Clark, 1999| Various tree species     | D     | NBA                | Nit: 2036, **2050**, 2078, 2152, **2180** |
| Curran et al., 2001 | Pine                     | D     | DRFL              | Nit: **1496**, 680, 648, **2352**, **2172** |
|                     |                          |       |                   | Pro: **1496**, **1008**, **1800**, **2352**, 1308 |
|                     |                          |       |                   | Chl: **622**, 2172, **1496**, **418**, **2352** |
|                     |                          |       |                   | Wat: **1936**, **1216**, **2172**, **1182**, **1920** |
| Serrano et al., 2002| Mediterranean vegetation | CA    | DRFL              | Nit: 1609, **1748**, 480 |
Table 7.7: The cross-validated coefficient of determination ($R^2$) and the cross-validated rmse between three types of spectral data (resampled to HyMap wavebands) and three leaf chemicals (year 1: $n=39$, year 3: $n=39$, year 1+3: $n=78$); DWRS: first-difference of reflectance for chlorophyll and water absorption features and first-difference of water removed spectra for protein absorption features; RFL: reflectance for all absorption features; BDN: band-depth normalized reflectance spectra for all absorption features; bands: number of wavebands selected by SMR; unit of relative rmse: % of mean

<table>
<thead>
<tr>
<th></th>
<th>DWRS</th>
<th></th>
<th>RFL</th>
<th></th>
<th>BDN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>rel. rmse</td>
<td>Bands</td>
<td>$R^2$</td>
<td>rel. rmse</td>
<td>Bands</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_{AB}$</td>
<td>0.81</td>
<td>7.1</td>
<td>3</td>
<td>0.83</td>
<td>6.4</td>
<td>3</td>
</tr>
<tr>
<td>$C_W$</td>
<td>0.60</td>
<td>4.5</td>
<td>3</td>
<td>0.53</td>
<td>4.9</td>
<td>6</td>
</tr>
<tr>
<td>$C_N$</td>
<td>0.63</td>
<td>6.1</td>
<td>6</td>
<td>0.54</td>
<td>6.6</td>
<td>3</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_{AB}$</td>
<td>0.63</td>
<td>9.7</td>
<td>3</td>
<td>0.58</td>
<td>10.3</td>
<td>3</td>
</tr>
<tr>
<td>$C_W$</td>
<td>0.63</td>
<td>5.1</td>
<td>7</td>
<td>0.39</td>
<td>6.5</td>
<td>3</td>
</tr>
<tr>
<td>$C_N$</td>
<td>0.72</td>
<td>4.7</td>
<td>5</td>
<td>0.64</td>
<td>5.2</td>
<td>4</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1+3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_{AB}$</td>
<td>0.81</td>
<td>8.7</td>
<td>3</td>
<td>0.80</td>
<td>9.0</td>
<td>3</td>
</tr>
<tr>
<td>$C_W$</td>
<td>0.77</td>
<td>5.9</td>
<td>3</td>
<td>0.75</td>
<td>6.2</td>
<td>5</td>
</tr>
<tr>
<td>$C_N$</td>
<td>0.52</td>
<td>6.7</td>
<td>4</td>
<td>0.49</td>
<td>6.8</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 7.8: Wavebands selected by SMR using three sets of spectral data (resampled to HyMap wavebands) and three leaf chemicals. DWRS: first-difference of reflectance for chlorophyll and water absorption features and first-difference of water removed spectra for protein absorption features; RFL: reflectance for all absorption features; BDN: band-depth normalized reflectance spectra for all absorption features; Directly related: ± 30 nm of the chemical of interest; Indirectly related: ± 30 nm of a chemical with which the chemical of interest was correlated (approach adapted from Curran et al., 2001)

<table>
<thead>
<tr>
<th>Spectral data</th>
<th>Leaf chemical</th>
<th>Total number of wavebands selected</th>
<th>Directly related</th>
<th>Indirectly related</th>
<th>Unattributed</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWRS</td>
<td>$C_{AB}$</td>
<td>9 (100)</td>
<td>1 (11)</td>
<td>2 (22)</td>
<td>6 (67)</td>
</tr>
<tr>
<td></td>
<td>$C_W$</td>
<td>13 (100)</td>
<td>3 (23)</td>
<td>6 (46)</td>
<td>4 (31)</td>
</tr>
<tr>
<td></td>
<td>$C_N$</td>
<td>15 (100)</td>
<td>9 (60)</td>
<td>2 (13)</td>
<td>4 (27)</td>
</tr>
<tr>
<td>RFL</td>
<td>$C_{AB}$</td>
<td>9 (100)</td>
<td>1 (11)</td>
<td>1 (11)</td>
<td>7 (78)</td>
</tr>
<tr>
<td></td>
<td>$C_W$</td>
<td>14 (100)</td>
<td>0 (0)</td>
<td>8 (57)</td>
<td>6 (43)</td>
</tr>
<tr>
<td></td>
<td>$C_N$</td>
<td>12 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>BDN</td>
<td>$C_{AB}$</td>
<td>14 (100)</td>
<td>2 (14)</td>
<td>1 (7)</td>
<td>11 (79)</td>
</tr>
<tr>
<td></td>
<td>$C_W$</td>
<td>9 (100)</td>
<td>0 (0)</td>
<td>1 (11)</td>
<td>8 (89)</td>
</tr>
<tr>
<td></td>
<td>$C_N$</td>
<td>10 (100)</td>
<td>5 (50)</td>
<td>0 (0)</td>
<td>5 (50)</td>
</tr>
</tbody>
</table>
7.4.3 Comparing different sets of spectral data

Three methods were compared on laboratory reflectance data. All three methodologies used SMR to develop relationships between chemical concentrations and a measure of reflectance in HyMap bands. The methods differed only in the spectral data used: First derivative of water removed spectra (DWRS), reflectance spectra (RFL), and band-depth normalised spectra (BDN). The spectra had been generated from the same five absorption features previously described. Between three and seven wavebands were selected with no marked differences in the accuracy of estimating chemical concentrations between the three data sets (Table 7.7). For year 1+3 samples, $R^2$ were uniformly high (>0.75) between all three forms of data and $C_{AB}$ and $C_W$. $R^2$ were moderate (about 0.50) between all three forms of data and $C_N$ for year 1+3 samples, but considerably larger between DWRS data and $C_N$ than between RFL data or BDN data and $C_N$ for year 1 and year 3 samples. There was a considerable difference between the three spectral data sets in the relevance of the selected wavebands with respect to known absorption features (Table 7.8). The total number of wavebands within ± 30 nm of the absorption feature of a chemical of interest (directly related) was largest for DWRS spectra and lowest for RFL spectra. For $C_N$, 73 percent of the selected wavebands were directly or indirectly related with DWRS data, but only 50 percent with BDN data and zero percent with RFL data. For $C_{AB}$, between 21 percent and 33 percent of the selected wavebands were directly or indirectly related with the three data sets. For $C_W$, the number of directly or indirectly related wavebands was 69 percent, 57 percent, and 11 percent for DWRS data, RFL data, and BDN data, respectively. The proportion of directly related wavelengths selected was considerably higher and the number of unattributed wavelengths selected was noticeably lower for DWRS data than for both other types of spectra.

7.4.4 Application of regression models to an independent data set

To determine the stability of the models, their performance with a completely different and independent validation data set than the calibration data (Gerolstein data set) was tested. The validation set (Idarwald data set) was a new set of samples measured under the same conditions as the calibration samples, however, from a different geographical site and date. To do so, relationships developed on the Gerolstein data set (n=78) were applied to the spectra of the Idarwald data set (n=30). The results were generally disappointing (Table 7.9). Coefficients of determination ($R^2$) between estimated and measured $C_{AB}$ reached 0.70 with F-data and did not considerably decrease with FH-data and FHN-data. However, the relative rmse between measured and estimated $C_{AB}$ values amounted from 25 percent up to 50 percent of the mean. For the other chemical attributes, though the values of relative rmse were comparatively low, values of $R^2$ were generally very low. A possible explanation for the large errors could be the time of sample collection. Samples used for calibration were collected in summer whereas collection of samples used for validation took place in spring.
7.5 Estimating chemical forest attributes from HyMap reflectance data

### 7.5.1 Scaling-up from leaf level to canopy level

Predictions of chlorophyll or other biochemical attributes from airborne or satellite canopy reflectance can be achieved through four different methodologies (Zarco-Tejada et al., 2001):

1. Establishing directly a statistical relationship between ground-measured chemical data and canopy-measured reflectance
2. Applying a leaf-level relationship derived between spectral indices and the chemical concentration to canopy-measured reflectance
3. Scaling up the leaf-level relationships based on spectral indices related to chemical attributes through models of canopy reflectance
4. Inverting the observed canopy reflectance through a canopy reflectance model coupled with a leaf model to estimate the chemical content
Methods 3 and 4 were not considered for estimating foliar chemicals in this work because the available forest reflectance model (INFORM), although it well simulated the variation of forest reflectance with changing canopy structure (section 6.3), did not prove to be very confident in modeling subtle changes of reflectance with change in chemical content in the visible domain and in parts of the mid-IR domain. Method 2 was not further considered as the application of leaf-level relationships developed on one site already showed relatively poor results when applied on another site and no improvement was expected when applying it to the canopy-level. Consequently, it was decided to choose method 1 for predicting chlorophyll, nitrogen and water concentrations from HyMap reflectance data.

To correlate chemical data with reflectances measured by the HyMap imaging spectrometer, chemical data that had been obtained on a leaf level had to be aggregated to stand level. Three different types of aggregation were performed: a) stand mean of year 1 needles, b) stand mean of year 3 needles, and c) stand mean of year 1 and year 3 needles. Canopy reflectance spectra extracted from the HyMap imagery (section 3.3.5) were processed in exactly the same way as the leaf reflectance data as described in section 7.2.1.

### 7.5.2 Comparing different sets of spectral data

Five methods were compared to develop relationships between chemical concentrations and HyMap reflectance (Table 7.10). Most often, one to three wavebands were selected. Relatively often, no regression model could be established. Differences in the accuracy of estimating attributes between the five types of derivatives were more pronounced than with laboratory data. None of the methods was capable in finding reasonable predictive equations to determine water concentration. Chlorophyll concentration ($C_{AB}$) for year 1+3 sample sets was well estimated by all methods except reflectance spectra (RFL). The best predictive model to estimate $C_{AB}$ was achieved by band-depth normalisation (BDN) with a cross-validated coefficient of determination ($R^2$) of 0.90 and a cross-validated relative rmse of 2.8 percent. Continuum-removed spectra (CR) performed less accurately (relative rmse: 4 percent), but used only 2 instead of 3 wavebands in the model. Nitrogen concentration was moderately well estimated by first-difference of water removed spectra (DWRS), CR and first-difference of reflectance spectra (DRFL) using three wavebands with a relative rmse below 5 percent. For year 1 and year 3 just one or two techniques showed good results while the other techniques performed poorly. The results are comparable to accuracies obtained in other studies. For instance, Mutanga & Skidmore (2004) mapped nitrogen concentrations in savanna grass with an rms error of 8.3 percent of the mean observed nitrogen concentration using continuum-removal.
Table 7.10: The cross-validated coefficient of determination ($R^2$) and the cross-validated rmse between five types of spectral data (HyMap imagery) and three leaf chemicals (n=13); relative rmse: % of mean; bands: number of wavebands selected by SMR; ***: $p_{in}=0.05$, $p_{out}=0.10$; **: $p_{in}=0.10$, $p_{out}=0.15$; *: $p_{in}=0.15$, $p_{out}=0.20$; units of rmse: % dry matter (C_W, C_N), mg g$^{-1}$ dm (C_AB); unit of relative rmse: % of mean

<table>
<thead>
<tr>
<th></th>
<th>DWRS R^2</th>
<th>Rmse (Rel. rmse)</th>
<th>Bands</th>
<th>R^2</th>
<th>Rmse (Rel. rmse)</th>
<th>Bands</th>
<th>R^2</th>
<th>Rmse (Rel. rmse)</th>
<th>Bands</th>
<th>R^2</th>
<th>Rmse (Rel. rmse)</th>
<th>Bands</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAB</td>
<td>0.67***</td>
<td>0.18 (6.4)</td>
<td>3</td>
<td>0.88**</td>
<td>0.10 (3.5)</td>
<td>4</td>
<td>0.44***</td>
<td>0.22 (7.8)</td>
<td>1</td>
<td>0.49***</td>
<td>0.21 (7.4)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_W</td>
<td>0.17***</td>
<td>7.6 (5.1)</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_N</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_AB</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.37***</td>
<td>0.30 (8.3)</td>
<td>1</td>
<td>0.75***</td>
<td>0.19 (5.3)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_W</td>
<td>0.47***</td>
<td>6.2 (5.0)</td>
<td>1</td>
<td>0.49**</td>
<td>6.7 (5.4)</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.22**</td>
<td>7.6 (6.2)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_N</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.19***</td>
<td>0.069 (5.9)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_AB</td>
<td>0.72***</td>
<td>0.15 (4.7)</td>
<td>3</td>
<td>0.46**</td>
<td>0.22 (6.9)</td>
<td>3</td>
<td>0.90***</td>
<td>0.09 (2.8)</td>
<td>3</td>
<td>0.80***</td>
<td>0.13 (4.0)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_W</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.45*</td>
<td>6.7 (4.9)</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_N</td>
<td>0.52***</td>
<td>0.058 (4.9)</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.34***</td>
<td>0.065 (5.5)</td>
<td>1</td>
<td>0.56**</td>
<td>0.058 (4.9)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Year 1

Year 3

Year 1+3
CHAPTER 7 ESTIMATION OF FOREST CHEMICAL ATTRIBUTES

The analysis of the selected wavebands (Table 7.11) revealed that a smaller number of bands could be directly related to known absorption features compared to laboratory reflectance data (Table 7.8). The largest portion of directly related wavebands for estimating nitrogen concentration again showed DWRS (67 percent) followed by DRFL (57 percent) and CR (33 percent). The wavebands chosen by stepwise multiple regression to determine the other leaf chemicals could not be directly related to known absorption features. However, selected wavebands, although not attributable to known features, were often located in spectral regions that are known to be important to detect chlorophyll (Table 7.12). For instance, the predictive model to estimate total chlorophyll concentration found with CR spectra was based on wavebands located along the red edge and along the edge of the green peak. This was also the case with the regression models developed on leaf spectra. Again, maximum correlations

Table 7.11: Wavebands selected by SMR using five sets of spectral data (HyMap imagery) and three leaf chemicals. DRFL-DWRS: first derivative of reflectance for chlorophyll and water absorption features and first derivative of water removed spectra for protein absorption features; RFL: reflectance; BDN: band-depth normalized reflectance spectra; CR: continuum-removed spectra; DRFL: first derivative of reflectance; Directly related: ± 30 nm of the chemical of interest; Indirectly related: ± 30 nm of a chemical with which the chemical of interest was correlated (approach adapted from Curran et al., 2001)

<table>
<thead>
<tr>
<th>Spectral data</th>
<th>Chemical attribute</th>
<th>Total number of wavebands selected</th>
<th>Directly related</th>
<th>Indirectly related</th>
<th>Unattributed</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWRS</td>
<td>C_{AB}</td>
<td>6 (100)</td>
<td>0 (0)</td>
<td>2 (33)</td>
<td>4 (67)</td>
</tr>
<tr>
<td></td>
<td>C_{W}</td>
<td>2 (100)</td>
<td>0 (0)</td>
<td>1 (50)</td>
<td>1 (50)</td>
</tr>
<tr>
<td></td>
<td>C_{N}</td>
<td>3 (100)</td>
<td>2 (67)</td>
<td>0 (0)</td>
<td>1 (33)</td>
</tr>
<tr>
<td></td>
<td>C_{AB}</td>
<td>7 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>7 (100)</td>
</tr>
<tr>
<td>RFL</td>
<td>C_{W}</td>
<td>6 (100)</td>
<td>0 (0)</td>
<td>4 (67)</td>
<td>2 (33)</td>
</tr>
<tr>
<td></td>
<td>C_{N}</td>
<td>0 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>C_{AB}</td>
<td>5 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>5 (100)</td>
</tr>
<tr>
<td>BDN</td>
<td>C_{W}</td>
<td>0 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>C_{N}</td>
<td>2 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (100)</td>
</tr>
<tr>
<td></td>
<td>C_{AB}</td>
<td>6 (100)</td>
<td>0 (0)</td>
<td>1 (17)</td>
<td>5 (83)</td>
</tr>
<tr>
<td>CR</td>
<td>C_{W}</td>
<td>1 (100)</td>
<td>0 (0)</td>
<td>1 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>C_{N}</td>
<td>6 (100)</td>
<td>2 (33)</td>
<td>0 (0)</td>
<td>4 (67)</td>
</tr>
<tr>
<td></td>
<td>C_{AB}</td>
<td>5 (100)</td>
<td>0 (0)</td>
<td>1 (20)</td>
<td>4 (80)</td>
</tr>
<tr>
<td>DRFL</td>
<td>C_{W}</td>
<td>2 (100)</td>
<td>0 (0)</td>
<td>1 (50)</td>
<td>1 (50)</td>
</tr>
<tr>
<td></td>
<td>C_{N}</td>
<td>7 (100)</td>
<td>4 (57)</td>
<td>1 (14)</td>
<td>2 (29)</td>
</tr>
</tbody>
</table>
occurred on the edge of absorption features. These findings are in full agreement with the results of other studies where differences in reflectance between healthy and stressed vegetation due to changes in pigment levels have been detected in the green peak and along the red edge (Rock et al., 1988; Gitelson & Merzlyak, 1996; Haboudane et al., 2002).

### 7.5.3 Maps of leaf chlorophyll concentration and leaf nitrogen concentration

Predictive models developed using CR spectra were applied on the HyMap image to compute maps of $C_{AB}$ and $C_N$. Only image pixels representing conifer forests were considered. The following predictive models between continuum-removed spectra (CR) and leaf chlorophyll concentration and leaf nitrogen concentration were applied to the HyMap image to derive the respective maps of foliar concentrations:

\[
C_{ab}^{est} = -41.7813 \cdot CR_{738} + 16.5392 \cdot CR_{585} + 33.3041
\]

where $C_{ab}^{est}$ is the estimated chlorophyll a+b concentration and $CR_{wavelength}$ denotes the continuum-removed value of reflectance in the specified waveband (in nm).

\[
C_{n}^{est} = -5.3535 \cdot CR_{738} + 2.7992 \cdot CR_{1157} - 8.9829 \cdot CR_{2373} + 12.724
\]

where $C_{n}^{est}$ is the estimated nitrogen concentration.

<table>
<thead>
<tr>
<th>Leaf chemical</th>
<th>Year</th>
<th>Selected wavebands [nm] and interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorophyll a+b</td>
<td>1</td>
<td>1157: Edge of water absorption band</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>738: Red edge, 753: Red edge, 2357: Protein at 2350</td>
</tr>
<tr>
<td></td>
<td>1+3</td>
<td>738: Red edge, 585: Edge of chlorophyll absorption feature</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Water</td>
<td>3</td>
<td>2341: Protein at 2350</td>
</tr>
<tr>
<td></td>
<td>1+3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>738: Red edge, 2373: Protein at 2350</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>3</td>
<td>738: Red edge</td>
</tr>
<tr>
<td></td>
<td>1+3</td>
<td>738: Red edge, 1157: Edge of water absorption band, 2373: Protein at 2350</td>
</tr>
</tbody>
</table>
The scatter plots (Figure 7.5) show linear relationships between measured and estimated values of $C_{AB}$ and measured and estimated values of $C_N$. The points in the scatter diagrams are close to the 1:1 line over the entire range of values indicating that no systematic underestimation or over estimation of the chemical concentrations occur.

In the chlorophyll map (Figure 7.6) and the nitrogen map (Figure 7.7), the observed values are in a reasonable range. $C_{AB}$ values range from 1 mg/g to 6 mg/g and $C_N$ values range from 1 % to 1.7 %. Most of the forest stands in both maps appear rather homogenous but also certain variations within the stands can be observed. From the interpretation of spatial patterns in both maps it can be noticed that $C_{AB}$ and $C_N$ are on the one hand interrelated to a certain extent but that on the other hand individual patterns exist.

### 7.5.4 Relations of foliar chemicals with other parameters

The maps of chemical forest attributes were integrated into the local Forest-GIS of Gerolstein. The Forest-GIS contained information about the tree species and the age structure on a stand base as well as the underlying geology. Statistical measures of chemicals were computed for zones of stand development phase and zones of geological epochs (Figure 7.8). Stand development phase is strongly related to stand age. Geological epoch stands for the rock types and the associated soil types that occur in the Gerolstein area and can be regarded as a measure of base saturation (Table 3.1) Mean values of chlorophyll concentration ($C_{AB}$) slightly increase from qualification towards maturation, but the large standard deviations indicate a substantial overlap between those classes. The same holds for zones of geological
epochs. At Gerolstein area, each epoch is roughly represented by a single rock type (section 3.1.2). Substrates from the Holocene and Pleistocene with high base saturation showed the largest mean chlorophyll concentrations. Again, large standard deviations pointed towards substantial overlap between the classes. To receive more distinct patterns, the chemical parameter maps were reclassified into three classes of low (class 1), medium (class 2) and high (class 3) concentration. For $C_{AB}$, the class limits were: $<2.8$ (class 1), 2.8-3.4 (class 2),
>3.4 (class 3). For nitrogen concentration, the class limits were: <1.15 (class 1), 1.15-1.25 (class 2), >1.25 (class 3). The reclassified parameter maps were then majority filtered using a 3x3 kernel. For each zone, the class majority (the class number that occurred most often) was computed. As can be seen from Figure 7.8, C_{AB} class 3 is the most frequent class over substrates with large base saturation whereas class 1 occurs most often over poorly base saturated soils over rocks from Buntsandstein (sands). For Lower Devonian and Mid
Devonian substrates, no relation between base saturation and $C_{AB}$ is evident. No relations were observed between nitrogen concentration and zones of stand development phase or zones of geological epoch.

An assessment concept for nitrogen and other biochemicals classifies values of nitrogen concentration in spruce needles (*Picea abies*) into five distinct classes (Hüttl, 1992): very low (< 1.3% dm), low (1.3-1.4% dm), medium (1.4-1.5% dm), high (1.5-1.7% dm), very high (> 1.7% dm). At Gerolstein test site, observed minimum nitrogen concentration (stand mean) is 1.08% of dry matter and the observed maximum is 1.31% dm. According to the classification scheme, 12 out of 13 stands have very low nitrogen concentrations. It is concluded that the large variation in base saturation present in the soils at Gerolstein area does is not equalled by corresponding variations in leaf nitrogen concentrations.

### 7.5.5 Accuracy requirements for ecological applications

Variations in foliar nitrogen occur between and within ecosystems. Foliar nitrogen in forests

![Statistical measure](image-url)

Figure 7.8: Statistical measures for three main phases and five main geological epochs that occur in the study area. Ordinate titles solely refer to mean and standard deviation.
across North America spans a range from 1 to 2.5 percent, as shown by Yin (1993, cited from Schimel, 1995), with foliar nitrogen being broadly correlated with summertime temperatures. Within ecosystems, seasonal variability in leaf nitrogen concentration can be very high; a change from early season nitrogen contents approaching 3 percent to late summer values of 0.8 percent in a tallgrass prairie was reported (Hobbs et al., 1991, cited from Schimel, 1995).

According to Schimel (1995), the following accuracies are required from remote sensing in order to map large scale variations in foliar nitrogen and lignin: Accuracy of 0.5% (absolute) nitrogen is necessary to distinguish between ecosystems with differences in nitrogen large enough to affect photosynthesis. Accuracy of 5% lignin concentration is needed to detect between-system gradients. Higher resolution, 1% lignin, is required to detect changes within a single forest type in order to map changes relevant to decomposition and nutrient cycling. These requirements are based on spatial gradients in nitrogen of 1-3.5 percent that occur between ecosystems. Spatial variations in spruce forest at Gerolstein site are considerably smaller: Stand mean values of C\textsubscript{N} (appendix A2) range from 1.08 percent to 1.31 percent. This range of C\textsubscript{N} (0.23 percent nitrogen) covers just about one tenth of what has been observed between ecosystems (2.5 percent nitrogen). From the low variation of nitrogen concentration present in Norway spruce needle at Gerolstein area no statement about the spatial variation of photosynthesis can be made. To increase variation in foliar nitrogen, the study has to be extended towards other ecosystems, such as beech forest or grassland.

Kokaly & Clark (1999) specified the required signal-to-noise ratio (SNR) to meet the accuracy requirements of Schimel (1995) as follows: A high SNR of approximately 700 is required to achieve an rmse of 0.5% absolute nitrogen. For RMS error of 5% absolute lignin concentration, the S/N must be only 100. However, the required SNR calculated from the dry leaves measured in the laboratory, with an average reflectance level of 40%, must raise by a factor of 8 for AVIRIS data over vegetation canopies, resulting in requirements for SNR of 800, 4400 and 5600 for lignin, cellulose, and nitrogen, respectively. They observed AVIRIS signal-to-noise ratios of about 360 at 2.10 µm. Their results show further, that it is possible to average two or more channels (bandpasses up to 78 nm) of the AVIRIS data, helping to compensate the lower SNR of the imaging spectrometer relative to laboratory measurements.

The SNR of the HyMap sensor, as specified by Cocks et al. (1998) based on ground measurement of a target with 50 percent reflectance, amount to 1600 at 0.55 µm, 1500 at 0.80 µm, and 900 at 2.10 µm. For a conifer forest target, the obtained SNR add up to 160, 600, and 90 at 0.55 µm (5 percent reflectance), 0.80 µm (20 percent reflectance), and 2.10 µm (5 percent reflectance), respectively. It is obvious, that the requirements as specified by Kokaly & Clark are not fulfilled by the HyMap sensor. Nonetheless, with an error of about 0.06 percent nitrogen (of dry matter) it was shown in the present research that an accurate determination of C\textsubscript{N} is possible. This can be attributed to the fact, that the required SNR, as
specified by Kokaly & Clark (1999) are based on absorption features in the mid-IR domain whereas the most important wavebands used in equation 7.2 are located in the red edge. Detection of variations in chlorophyll and thus, nitrogen along the red edge requires much lower SNRs.
Chapter 8  Conclusions and future perspectives

Research overview

It has been the overall aim of this research work to assess the potential of hyperspectral remote sensing data for the determination of forest attributes relevant to forest ecosystem simulation modeling and forest inventory purposes. The research contributes to the fundamental concern of remote sensing; that is, to evaluate to what extent remotely sensed radiation can in fact provide useful information for a specific application and to develop the appropriate retrieval techniques (Verstraete et al., 1996).

The logic applied to the research followed three distinct phases:

1. Attributes that describe the properties of the forest canopy and that are potentially open to remote sensing were identified, appropriate methods for their retrieval were implemented and field, laboratory and image data were acquired over a number of forest plots.
2. The forest plots served as calibration and validation base for deriving relationships between forest attributes and remotely sensed data using image classification techniques, empirical-statistical approaches and physically-based methods.
3. The remotely sensed data were used to extrapolate plot estimates of forest attributes over the entire study area to produce parameter maps; the resulting spatial patterns were analysed with respect to other environmental factors; finally, the accuracies to predict the attributes were compared to those needed for applications in forestry.

Selection of forest attributes

Among the variety of attributes that characterise forest canopies, three requirements were identified to select suitable forest attributes:

i) the attribute has to be of interest to ecological or economical applications
ii) the attribute must directly or indirectly affect the transfer of radiation to allow the retrieval of reliable quantitative information on that attribute by remote sensing
iii) ground measurement of the attribute has to be feasible using widely accepted and operational methods in the field and laboratory. The attributes selected in this research work that affect ecological processes are leaf area index (LAI), leaf chlorophyll a+b concentration (C_AB), and leaf nitrogen concentration (C_N). Attributes that are of
concern to forest inventory and management are stem biomass (SBM), crown volume (VOL) and stand age.

Retrieval of forest attributes from hyperspectral data

A number of approaches for the determination of structural and chemical attributes from hyperspectral remote sensing have been applied to the collected data sets. Many of the methods to be found in the literature were

a) applied to broadband multispectral data
b) applied to vegetation canopies other than forests
c) reported to work on the leaf level or with modelled data
d) not validated with ground truth data or
e) not systematically compared to other methods.

Mapping of forest type

Age structure and species composition of forest stands are routinely collected in forest inventories every 10 years or longer to assess timber store and to optimise harvesting. Data of many spectral and spatial resolutions can be used for mapping forest type. In this study, the goal is to investigate whether tree species and age classes can be reliably derived from hyperspectral remote sensing data through image classification and to compare the overall information content of fine to broad spectral resolution and small to large spatial resolution data. The results of the image classification show that the overall accuracy and kappa coefficient using hyperspectral data are substantially larger than with multispectral data. These results indicate that hyperspectral data contain more information relevant to the mapping of forest type than broadband data. Remarkably, degradation of the high spatial resolution to coarser image pixels considerably increased the classification accuracies. In hyperspatial data individual trees comprise several pixels, each pixel may represent one of several classes (shadow or sunlit, young foliage or old), and the larger class variance may reduce class discrimination.

For mapping of forest types in Germany where currently the vast majority of the forests are one-age-cohort plantations, a moderate spatial resolution (30 m) optical sensor would be preferable. One-age-cohort management is not sustainable from an ecological point of view and has been abandoned by the forest administrations of all German federal states (Häusler & Scherer-Lorenzen, 2001). It is currently being replaced by alternative management strategies that in the future will increasingly produce forest stands with rather heterogeneous species composition and age structure. From a remote sensing perspective, under these conditions a hyperspatial sensor would be preferable that could detect such small scale variations to avoid the problem of mixed pixels. A hyperspectral would be preferred over a multispectral system to make use of narrow spectral features for discriminating different forest types.
Required classification accuracies from remote sensing, as claimed by Foody (2002), are not fulfilled by most of the datasets except hyperspectral data with low spatial resolution. One possible explanation would be that the classification problem encountered in this study is not trivial. Forest classes, such as clearings have spectral reflectance characteristics that differ very much from the closed canopy stands studied here. Those clearcuts are relatively easy to discriminate from the other forest types and including them in this study would probably have increased the classification accuracy.

Improvement of forest type mapping is often attempted by adding ancillary data into the classification process. The influence of stem density information proves to have a limited effect on classification accuracy. The major disadvantage of this approach is the restricted availability of these data. Other methods overcome traditional pixel-based image analysis and make use of textural information. Using geostatistical parameters derived from the image data itself as demonstrated by Buddenbaum et al. (2005) is another innovative approach to improve classification results.

**Determination of leaf area index**

Leaf area index (LAI) plays an important role in the quantification of ecosystem process, e.g. absorption of photosynthetically active radiation, canopy conductance and interception of precipitation. LAI, as other attributes related to canopy cover, affects the remotely sensed signal in sparse to medium dense canopies significantly. The goal is to compare empirical and physical approaches to determine LAI and to identify the most suitable hyperspectral and multispectral bands for both methods. The main results obtained from the research are summarised in Table 8.1. Determination of LAI through linear regression between remotely sensed vegetation indices (VIs) and ground measured LAI can be performed best with a narrowband perpendicular vegetation index (PVI) based on wavebands at 1088 nm and 1148 nm ($R^2 = 0.77$, rmse = 0.54 m$^2$ m$^{-2}$). LAI determination with NDVI and other broadband ratio vegetation indices based data resampled to Landsat TM wavebands is less successful due to saturation effects, influence of background and understorey vegetation. These effects are partly compensated when soil-line related broadband VIs (e.g. PVI) are used ($R^2 = 0.59$, rmse = 0.84 m$^2$ m$^{-2}$). However, problems with old stands are still present that can be ascribed to shadow effects and a relatively dark background in the near-IR.

The physical approach inverts a forest reflectance model using a neural network approach. With wavebands at 837 nm and 1148 nm the obtained accuracy of the retrieved LAI map amounts to an rmse of 0.58 m$^2$ m$^{-2}$. With HyMap data resampled to Landsat TM spectral bands and using two bands at 840 nm and 1650 nm, an rmse of 0.66 (relative rmse = 20 % of the mean) is achieved. Comparing the performance of both, empirical and physically-based approaches for LAI estimation from hyperspectral data it is concluded that both methods work
equally well. When broadband multispectral data is used, LAI is more accurately predicted by the reflectance model inversion. This is a clear advantage of the physically-based approach over the empirical models. The physical approach explicitly takes into account canopy architecture, multiple scattering and signal contributions of the background and understorey vegetation. Quite different from empirical prediction functions that are generally limited to the local conditions at a certain point of time and a specific sensor type, the calibrated reflectance model can be applied more easily to different optical remote sensing data acquired at landscape to regional scales.

Analysis of the relations between best narrow band PVIs and LAI reveals that the wavebands forming the vegetation indices are either located at the shoulder of a water absorption feature (1088 nm) or at the absorption minimum (1148 nm). The same waveband (1148 nm) is employed to invert the reflectance model and the other waveband is situated on the near-IR plateau (837 nm). Obviously, in both, the empirical and the physical approach, the total amount of canopy water is mapped by the 1148 nm waveband. As the leaf water content is considered to be relatively constant in the study area, variations in LAI are retrieved.

Attributes related to aboveground biomass can not be accurately retrieved by both empirical and physically-based approaches. Poor relations are found between various vegetation indices and stem biomass. However, a synthetic variable is derived from measured structural (biomass related) attributes (stem density, stem diameter, canopy height) through factor analysis. This variable is interpreted as measure of crown volume (VOL) and it is linearly related to a PVI based on wavebands at 885 and 948 nm ($R^2 = 0.79$, rmse = 0.52 m$^2$ m$^{-2}$). Reflectance model inversion considerably underestimates stem density compared to ground data for dense forest stands (rmse 55 % of mean). To retrieve attributes related to aboveground biomass different remote sensing data and techniques have to be employed. For instance, lidar systems can very accurately measure the height of tree canopies and allow determining timber store.

<table>
<thead>
<tr>
<th>Method</th>
<th>Data/Index (wavebands)</th>
<th>$R^2$</th>
<th>Rmse [m$^2$ m$^{-2}$]</th>
<th>Relative rmse [% of mean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical</td>
<td>Hyperspectral PVI (1088 nm, 1148 nm)</td>
<td>0.77</td>
<td>0.54</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Broadband PVI (670 nm, 840 nm)</td>
<td>0.59</td>
<td>0.84</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Broadband NDVI (670 nm, 840 nm)</td>
<td>0.47</td>
<td>1.06</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Hyperspectral (837 nm, 1148 nm)</td>
<td>0.73</td>
<td>0.58</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Broadband (670 nm, 840 nm)</td>
<td>0.69</td>
<td>0.85</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Broadband (840 nm, 1650 nm)</td>
<td>0.73</td>
<td>0.66</td>
<td>21</td>
</tr>
<tr>
<td>Physically-based</td>
<td>Broadband PVI (670 nm, 840 nm)</td>
<td>0.59</td>
<td>0.84</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Broadband NDVI (670 nm, 840 nm)</td>
<td>0.47</td>
<td>1.06</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Hyperspectral (837 nm, 1148 nm)</td>
<td>0.73</td>
<td>0.58</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Broadband (670 nm, 840 nm)</td>
<td>0.69</td>
<td>0.85</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Broadband (840 nm, 1650 nm)</td>
<td>0.73</td>
<td>0.66</td>
<td>21</td>
</tr>
</tbody>
</table>
CHAPTER 8 CONCLUSIONS AND FUTURE PERSPECTIVES

Determination of chlorophyll, water and nitrogen concentration

The study on chemical forest attributes evaluated the information content of HyMap data for the estimation of chlorophyll (C<sub>AB</sub>) water (C<sub>W</sub>) and nitrogen (C<sub>N</sub>) concentration. The aims were to find predictive models to estimate chemical concentrations and to systematically compare different types of derivatives of spectral reflectance regarding the accuracy of prediction. The results of the analysis (Table 8.2) show that C<sub>AB</sub> can be well estimated from laboratory and canopy reflectance data. The best predictive model to estimate C<sub>AB</sub> at the laboratory level yielded a cross-validated R<sup>2</sup> of 0.81 and a relative rmse of 8.7 % (n =78). The best model to estimate C<sub>AB</sub> at the canopy level was achieved using band depth normalised spectra with a cross-validated R<sup>2</sup> of 0.90 and a relative rmse of 2.8 %. However, the small sample size at the canopy level (n = 13) is critical. The promising results that are obtained for chlorophyll may be explained by two reasons. First, chlorophyll is a strong absorber of radiation and variations of concentration have a large influence on reflectance. Second, chlorophyll concentration shows the largest variation of all chemicals which may explain the strong statistical correlation between chlorophyll concentration and reflectances. The same factors operating conversely may explain the weaker relations between reflectance and the concentration of nitrogen (R<sup>2</sup> = 0.57, rel. rmse = 4.6 %). Protein, one of the main nitrogen bearing leaf constituents, is a weak absorber of radiation and had a low variation between samples. Water concentration is well estimated from laboratory reflectance data, but not from noisy laboratory spectra and canopy spectra. This seems to be surprising as leaf water is a strong absorber of radiation. However, the low variation between samples could be the reason for the weak correlation between water concentration and noisy or canopy reflectance data.

The analysis of the selected wavebands reveals that for canopy spectra a smaller number of bands can be directly related to known absorption features compared to laboratory reflectance data.

Table 8.2: Determination of chemical forest attributes using laboratory and HyMap data and three types of spectral derivatives; relative rmse: % of mean; percentage of attributable bands: directly or indirectly related to the variable of interest

<table>
<thead>
<tr>
<th>Data</th>
<th>DWRS</th>
<th>RFL</th>
<th>BDN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Rel. rmse</td>
<td>Perc. of attrib. bands</td>
</tr>
<tr>
<td>Laboratory resampled to HyMap (n=78)</td>
<td>C&lt;sub&gt;AB&lt;/sub&gt;</td>
<td>0.81</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>C&lt;sub&gt;W&lt;/sub&gt;</td>
<td>0.77</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>C&lt;sub&gt;N&lt;/sub&gt;</td>
<td>0.52</td>
<td>6.7</td>
</tr>
<tr>
<td>HyMap image (n=13)</td>
<td>C&lt;sub&gt;AB&lt;/sub&gt;</td>
<td>0.72</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>C&lt;sub&gt;W&lt;/sub&gt;</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>C&lt;sub&gt;N&lt;/sub&gt;</td>
<td>0.52</td>
<td>4.9</td>
</tr>
</tbody>
</table>
data. The wavebands selected to estimate $C_{AB}$ with laboratory and HyMap spectra are typically located in the red edge region and along the edge of the chlorophyll absorption feature. This result coincides with a general understanding that the maximum correlations occur on the edge of absorption features as the central wavelengths become saturated. For $C_N$, the wavebands selected in the regression models with laboratory data are often related to the chlorophyll absorption feature, the red edge region and to known protein absorption features. However, the portion of selected wavebands attributable to known substance absorption features strongly depends on the type of derivative spectra used. A method called “derivative of water removed spectra” (DWRS) produces the largest percentage of wavebands directly or indirectly related to known absorption features with laboratory and canopy spectra. Using band depth normalised spectra (BDN) the portion of wavebands attributable to known absorption features was relatively large with laboratory data but very low with HyMap spectra. Possibly, DWRS better than BDN preserves its ability to enhance absorption features during the upscaling from leaf to canopy spectra.

Wavelengths close to protein features at 1730 nm, 2180 nm, and 2350 nm that have often been selected in the literature are also selected in this research. As in the present work, a combined selection of wavebands related to both chlorophyll and protein absorption features in a single model to estimate nitrogen concentration is also reported from other studies.

The ability to estimate chemical concentrations from reflectance spectra using a model derived from an independent sample set is fundamental for extrapolating results over large sample sets and for determining of the model’s long term stability. Here, two independent data sets of Norway spruce samples were used. Estimations of all chemicals had large errors and low $R^2$ values except chlorophyll which showed values of $R^2$ above 0.65 for noisy laboratory spectra.

A limited data quality was reported for early AVIRIS campaigns (Kupiec, 1994; Niemann & Goodenough, 2003) which did not allow an accurate mapping of foliar chemicals. Contrary, the results of this research indicate that the detection and mapping of variations in chlorophyll and nitrogen concentrations is possible through the use of HyMap data. Obviously, the signal-to-noise ratio of the HyMap sensor is large enough to allow detection of small reflectance variations. Apart from the quality of the raw data, a sound radiometrically calibration to retrieve surface reflectances is an essential requirement for successful mapping of quantitative forest attributes. It can be concluded that the radiometric processing applied to the data fulfils the requirements to allow an accurate mapping of structural and chemical forest attributes.

### Reflectance modeling

An essential requirement for model inversion is to ascertain that the reflectance signatures simulated by the model are consistent with the remotely sensed canopy spectra. If correctly
parameterised, the INFORM model simulates forest reflectance spectra comparable to those measured by an imaging spectrometer (HyMap). There is a good agreement between modelled and measured spectral reflectance in the near-IR wavebands and an acceptable agreement in the mid-IR. In the visible, however, model outputs and HyMap reflectances disagree. The disagreement in the visible domain can be explained by the used leaf reflectance model. A possible solution for the problem in this spectral domain could be to replace the LIBERTY by the PROSPECT model which was shown in other studies to work also with conifer needles. Perhaps better agreement between simulated and HyMap spectra in the visible could be obtained if chlorophyll content was changed within the typical range of variation. Also the performance of INFORM in the mid-IR region has to be further improved. The difference between simulated HyMap spectra in the mid-IR could be the reason for the poor estimation of stem density. Future studies will continue to clarify the relative importance of structural factors influencing the remote sensing measurements collected over forests.

A central question of the research on physically-based models was to study a new method to invert the model based on neural networks. The applied network with two input neurons and only one neuron in the hidden layer is of a very simple structure. Presumably, the two input neurons together establish a linear combination of two wavebands, similar to PVI. Next, the hidden neuron transforms the intermediate result in non-linear way, after which three output neurons compute the output to approximate the requested output variables. Finally from three output variables canopy LAI is computed. The assumption that the neural network approach works in comparable manner to PVI is supported by the fact that similar wavebands are used in both, PVI-based and neural network based approaches and that equal accuracies are retrieved by both methods.

Spatial patterns of parameter maps

After the best empirical model was applied to the HyMap image to derive estimates of LAI for each pixel, pixel-based LAI was aggregated to mean stand values. After a peak LAI at an age of 20, mean LAI is decreasing with increasing stand age. The observed age course is a result of both natural circumstances and actual management practices. Concurrently, regeneration of leaf area index subsequent to thinning measures is also determined by the water and nutrient supply of a particular site. Standard deviation of stand LAI is particularly large for the age class of 31-50 year old stands. The reason for the large variability of the 31-50 year old stands is that some stands of this age may be subjected to thinning measures while others may remain undisturbed and thinned at a later stage. LAI can vary considerably within a single age class and thus, information on LAI can not simply be derived from an age class map.

Predictive models were applied on the HyMap image to compute maps of chlorophyll concentration and nitrogen concentration. Results of map overlay operations revealed
coherence between $C_{AB}$ and zones of stand development stage and between $C_{AB}$ and zones of soil type. Larger values of $C_{AB}$ occurred over soils with large base saturation (formed of basalt and colluvium) whereas poorly base saturated soils over sandstone showed lower $C_{AB}$ values. However, the large variation in base saturation present in the soils at Gerolstein area is not equalled by corresponding variations in leaf nitrogen concentrations. From the low variation of nitrogen concentration present in Norway spruce needle at Gerolstein area no statement about the spatial variation of photosynthesis can be made. To increase variation in foliar nitrogen, the study has to be extended towards other ecosystems, such as deciduous forest, vine yards and orchards or grassland.

Future perspectives

Based on the results attained in present investigation, diverse perspectives for improved remote sensing of forests emerge:

The methods developed to determine structural forest attributes have evolved considerably in the past few decades. Vegetation indices which are by far the most widely used method are computationally efficient, but difficult to interpret. Both field and modeling studies show that vegetation indices are difficult to interpret with respect to a single structural forest attribute. Vegetation indices will continue to play an important role in remote sensing of forests and their improvement over time is warranted (Asner et al., 2003).

The use of canopy reflectance models to retrieve forest attributes from remote sensing data is still at its beginning. Most studies so far focused on evaluating model performance. Many of the newest 3-D models can accurately simulate the spectral and angular reflectance properties of forests. Future forward modeling studies will improve the understanding of how structural and chemical attributes govern the reflectance of forest canopies (Asner et al., 2003). On the other hand, little research has been done to invert models with remotely sensed data and validate the retrieved forest attributes with ground truth data.

To improve mapping of forest attributes through inverse modeling, future efforts should focus on developing new inversion techniques and making use of additional data dimensions. Improved inversion strategies could, for instance, combine an optimal spectral sampling with algorithms to find powerful predictive models. Lately, genetic algorithms have been identified as a powerful tool for finding optimal set of spectral variables to be included into a model. The utilisation of multi-view angle data is considered a prime option for advanced quantitative analysis of forested areas. Currently, the CHRIS sensor on the PROBA platform provides hyperspectral data with 17-34 m spatial resolution collected in five different view angles. Using such data could probably improve mapping of forest structural attributes.

In the future, the users of remote sensing products will ask for a growing selection of higher-level and value-added products that are consistent in quality, reliably available and easy to
use. This requires that validation and quality considerations of data products will have to receive much more attention than in the past (Peddle et al., 2003).
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


KENNEDY, R. O., COHEN, W. B., TAKAO, G. 1997: Empirical methods to compensate for a
view-angle dependent brightness gradient in AVIRIS imagery. Remote Sensing of

KIMBALL, J. S., WHITE, M. A., RUNNING, S. W. 1997: Biome-BGC simulations of stand
hydrologic process for BOREAS. Journal of Geophysical Research, 102(D24): 29043-
29051.

KIMES, D. 1983: Dynamics of directional reflectance factor distribution for vegetation

KIMES, D. S., HOLBEN, B. N., NICKESON, J. E., MCKEE, W. A. 1996: Extracting forest age in a
Pacific Northwest forest from Thematic Mapper and topographic data. Remote Sensing of
Environment, 56: 133-140.

KIMES, D., GASTELU-ETCHEGORRY, J, ESTÈVE, P. 2002: Recovery of forest canopy
characteristics through inversion of a complex 3D model. Remote Sensing of


KOKALY, R. F. & CLARK, R. N. 1999: Spectroscopic determination of leaf biochemistry using
band-depth analysis of absorption features and stepwise multiple linear regression.

KOKALY, R. F. 2001: Investigating a physical basis for spectroscopic estimates of leaf

KÖTZ, B., SCHAEPMAN, M., MORSORF, F., BOWYER, P., ITTEN, K., ALLGÖWER, B. 2004:
Radiative transfer modeling within a heterogenous canopy for estimation of forest fire


sensing for the Earth sciences: Manual of remote sensing, 3rd ed., Vol. 3. New York,

System (SIPS) - Interactive visualization and analysis of imaging spectrometer data.

KUPIEC, J. A. & CURRAN, P. J. 1995: Decoupling effects of the canopy and foliar

Ph.D.-thesis.

KUUSK, A. 1998: Monitoring of vegetation parameters on large areas by the inversion of a

LAMB, D. W., STEYN-ROSS, M., SCHAEP, P., HANNA, M. M., SILVESTER, W., STEYN-ROSS,
A. 2002: Estimating leaf nitrogen concentration in ryegrass (Lolium spp.) pasture using
the chlorophyll red-edge: theoretical modeling and experimental observations.


REFERENCES


ROUSE, J.W., HAAS, R.H., SCHELL, J.A. 1974: Monitoring the vernal advancement of retrogradation of natural vegetation. Greenbelt, MD, USA.


Summary

Forest ecosystems cover large parts of the Earth’s surface and play an important role in the global carbon budget. An understanding of their function requires determining their properties in space and time. The ability to remotely sense forest canopy properties would offer great potential for applications at regional to global scales such as environmental change assessment, forest inventory and management, and monitoring of forest fire and disease. A relatively new concept, hyperspectral remote sensing, may provide more information about the properties of forest canopies than conventional broadband sensor types, such as Landsat or SPOT.

Previous research studies in remote sensing attempted to estimate forest attributes with varying success. A lot of research has been done on the estimation of forest structural attributes, in particular leaf area index (LAI) from remote sensing data within about the last 15 years. Most of the studies on forests used broadband multispectral data, but very recently a few authors used hyperspectral remote sensing data to assess forest LAI. Few studies have reported on the application of multi- and hyperspectral data in mapping forest species and forest age classes. Laboratory studies on dry and ground leaves showed strong relations between leaf reflectance and foliar chemicals, but results on fresh whole leaves were generally poorer. From intensive research activities during the last decade it was concluded that a reliable estimation of canopy chemical concentrations from remote sensors needs an improvement of data quality and analysing techniques.

The aim of this research was to investigate the use of hyperspectral remote sensing data to determine important forest attributes. The work was conducted in forest plots (Picea abies L. Karst.) in Eifel and Hunsrück mountains, Germany. It focused on three broad types of forest attributes: cover type, structural attributes and chemical attributes. As such information is not explicitly recorded by a sensor it has to be extracted from remote sensing data for which a number of methodological approaches are generally available. The main focus of this work was to use and modify existing methods and to develop new approaches to successfully determine the forest attributes of interest. Major attention was paid to systematically compare different methods and data types.

The study on forest cover type compared the classification performance of four different remote sensing datasets of different spectral and spatial resolution to map forest species and age classes. To generate synthetic, multispectral broadband data, the HyMap image was spectrally resampled to Landsat TM wavebands and spatially degraded to 30 meter pixels. The results of the image classification showed that the overall accuracy and Kappa coefficient using the HyMap datasets were substantially larger than using the TM datasets. These results confirm that hyperspectral data contain more information relevant to the mapping of forest type than broadband multispectral data. Interestingly, degradation of the high spatial resolution to coarser image pixels considerably increased the classification accuracies.

The study on structural attributes compared empirical and physically-based approaches. In the empirical section linear predictive models between vegetation indices (VIs) derived from HyMap data and field measurements of structural forest stand attributes (n=39) were systematically evaluated. Ratio-based and soil-line related broadband VIs were calculated after HyMap reflectance had been spectrally resampled to Landsat TM channels.
Hyperspectral VIs involved all possible types of 2-band combinations of ratio VIs (RVI) and perpendicular VIs (PVI) and the red-edge inflection point (REIP) computed from two techniques. A PVI based on wavebands at 1088 nm and 1148 nm was linearly related to leaf area index (R$^2=0.77$, rmse=0.54 m$^2$m$^{-2}$ (17% of mean)). A similar index was linearly related to the crown volume that had been derived from measured structural attributes through factor analysis. The study demonstrates that for hyperspectral image data, linear regression models can be applied to quantify LAI and VOL with good accuracy. For broadband multispectral data, the accuracy was generally lower. Best hyperspectral VIs in relation with LAI were typically based on wavebands related to prominent water absorption features.

The physically-based approach used the invertible forest reflectance model (INFORM), a combination of well established sub-models FLIM, SAIL and LIBERTY. The model was inverted with HyMap data using a neural network approach. To identify the optimum 2-band spectral subset to be used in the inversion process, all 2-band combinations were systematically evaluated for model inversion. LAI field measurements from 39 forest stands were used to validate the LAI maps produced from HyMap imagery. Using two HyMap wavebands at 837 nm and 1148 nm the obtained accuracy of the LAI map amounts to an rmse of 0.58 m$^2$m$^{-2}$ (18% of mean). The approach could be shown to potentially work also on operational multispectral sensors. With HyMap data resampled to Landsat TM spectral bands and using two bands at 840 nm and 1650 nm, the rmse was 0.66 and relative rmse amounted to 21 %. In comparison to the empirical approach, it could be shown that the reflectance model inversion works equally well. In opposition to empirically derived prediction functions that are generally limited to the local conditions at a certain point in time and to a specified sensor type, the calibrated reflectance model can be applied more easily to different optical remote sensing data acquired over central European forests.

The study on chemical forest attributes evaluated the information content of HyMap data for the estimation of nitrogen, chlorophyll and water concentration. A number of needle samples (n=78) of Norway spruce were analysed for their total chlorophyll (C$_{AB}$), nitrogen (C$_N$) and water (C$_W$) concentrations. The research was divided into two areas: needle spectra measured in the laboratory and canopy spectra measured by the HyMap sensor. The objective was to find empirical models by stepwise multiple linear regression (SMR) analysis using certain absorption features to determine foliar chemicals.

Laboratory reflectance measurements were made for layer stacks of whole green needles. To test the transferability of the approach from the leaf to the canopy level, the effects of sensor bandwidth and sensor noise on the estimation of the chemical concentrations was investigated. SMR analysis showed that C$_{AB}$ was well estimated and just slightly affected by a decrease of the spectral resolution or adding of noise, C$_N$ was moderately well estimated and just slightly affected by a decrease of the spectral resolution or adding of noise and C$_W$ was well estimated but strongly affected by noise addition. The wavebands selected to estimate C$_{AB}$ were typically located in the red edge region and along the edge of the chlorophyll absorption feature. For C$_N$, the wavebands selected in the regression models were often related to the chlorophyll absorption feature, the red edge region and to known protein absorption features. Regression models to estimate C$_W$ often employed wavebands related to C$_{AB}$ or wavebands located at the edge of water absorption features. Scaling up from leaf to canopy level was achieved through directly establishing statistical relationships between ground-measured chemical data and canopy-measured reflectances. Five different types of derivatives of reflectance were compared in estimating foliar chemicals: reflectance, first-difference of reflectance, continuum-removed spectra, band-depth normalised spectra (BDN) and water removed spectra. The best predictive model to estimate C$_{AB}$ was achieved by BDN wit a cross-validated R$^2$ of 0.90 and a relative rmse of 2.8 percent (n=13). C$_N$ was moderately well estimated (relative rmse below 5 percent). Selected wavebands, although most often not attributable to known absorption features, were often located in spectral regions that are known to be important for chlorophyll detection (red edge, green peak). Predictive models were applied on the HyMap image to compute maps of chlorophyll concentration and nitrogen concentration. Results of map overlay operations revealed coherence between C$_{AB}$ and zones of stand development stage and between C$_{AB}$ and zones of soil type. Larger values of C$_{AB}$ occurred over soils with large base saturation (formed
of basalt and colluvium) whereas poorly base saturated soils over sandstone showed lower $C_{AB}$ values.

Finally, it can be stated that the hyperspectral remote sensing data generally contains more information relevant to the estimation of the forest attributes compared to multispectral data. Structural forest attributes can be determined with good accuracy from a hyperspectral sensor type like HyMap. Among the chemical attributes, chlorophyll concentration can be determined with good accuracy and nitrogen concentration with moderate accuracy. For future research, additional dimensions have to be taken into account, for instance through exploitation of multi-view angle data.
Zusammenfassung

Die fernerkundliche Erfassung struktureller und chemischer Forstattribute ist in der aktuellen Ökosystemforschung, der Entwicklung zukünftiger Inventarisierungsverfahren und für das Umweltmonitoring von hoher Relevanz. Bisherige Untersuchungen zur Quantifizierung von Forstattributen aus Fernerkundungsdaten, die zumeist auf der Auswertung multispektraler Satellitenbilder basierten, verliefen mit unterschiedlichem Erfolg.


Für die Bestimmung der strukturellen Forstattribute wurden empirisch-statistische und physikalisch-basierte Verfahren implementiert. Im empirischen Teil der Studie wird der Zusammenhang zwischen Vegetationsindices, die aus HyMap-Daten berechneten wurden, und im Gelände gemessenen Bestandsvariablen analysiert. Die Geländeerhebungen umfassen
die standardisierte Bestimmung wichtiger Strukturgrößen in 42 Fichtenbeständen, darunter Blattflächenindex, Bestockungsdichte, Kronenschlussgrad, Brusthöhen- und Bestandeshöhe. Der effektive Blattflächenindex (LAI) wurde über ein optisches Verfahren (LI-COR LAI-2000) bestimmt. Es werden unterschiedliche breit- und schmalbandige sowie „ratio“- und „soil line“-basierte Vegetationsindices und darüber hinaus Verfahren zur Bestimmung des Inflexionspunktes systematisch untersucht. Die ausgewählten breitbandigen Vegetationsindices (SR, NDVI, PVI, SAVI, TSAVI, u. a.) wurden zwar schon in Einzelstudien angewendet, bisher aber noch nicht einem systematischen Vergleich auf Wäldern unterzogen. Die schmalbandigen VIs umfassen sämtliche 2-Band Kombinationen hyperspektraler „Ratio“-VIs (RVI) und „Perpendicular“-VIs (PVI). Als GüteMaß zur Bewertung der unterschiedlichen Indizes dienen kreuzvalidiertes Bestimmungsmaß ($R^2$) und kreuzvalidierter rms Fehler („root mean squared error“). Best er Index ist ein hyperspektraler PVI basierend auf Spektralbändern bei 1088 nm und 1148 nm, der linear zum LAI in Beziehung gesetzt wird ($R^2 = 0.77$, $\text{rmse} = 0.54 \text{m}^2 \text{m}^{-2}$ (17% des Mittelwertes)). Ein ähnlicher Index zeigt einen engen Zusammenhang zum Kronenvolumen, das zuvor aus den im Gelände erhobenen Attributen mittels Faktorenanalyse bestimmt worden ist. Breitband VIs erzielen generell geringere Genauigkeiten. Zur Stammbiomasse konnten keine Beziehungen gefunden werden. Die Positionen der in den Indizes verwendeten Spektralbänder befinden sich in der Nähe prominenten Wasserabsorptionsbanden – ein Beweis dafür, dass der Differenzierung nach LAI maßgeblich der Bestandswassergehalt zu Grunde liegt. Die Bedeutung der Spektralbänder im mittleren Infrarot für die Charakterisierung von Wäldern ist auch durch andere Untersuchungen bestätigt worden. Ferner spiegelt die festgestellte Veränderung des LAI mit dem Bestandsalter deutlich die waldbaulichen Maßnahmen wider. Die Studie demonstriert, dass Hyperspektraldaten gut zur Quantifizierung wichtiger Forstattribute wie LAI und Kronenvolumen, aber nicht biomassebezogener Größen geeignet sind.

Der physikalisch-basierte Ansatz verwendet das Invertierbare Forstreflexionsmodell (INFORM), eine Kombination aus etablierten Untermodellen FLIM, SAIL und LIBERTY. Nachdem das Modell mit Attributdaten aus Geländeerhebungen parametrisiert worden ist, wird es auf HyMap-Daten mittels Künstlicher Neuronaler Netze (KNN) invertiert. Zur Identifizierung geeigneter Spektralkanäle werden bei der Inversion alle 2-Band Kombinationen systematisch untersucht. Die aus der Modellinversion resultierenden Parameterkarten werden mit den Geländedaten validiert. Mit zwei Spektralbändern bei 837 nm und 1148 nm kann für LAI eine Fehlerr genauigkeit von 0.58 m² m⁻² (18% vom Mittelwert) erzielt werden. Auf spektral reduzierten Daten (Landsat TM) ergibt sich unter Verwendung der Spektralbänder bei 840 nm und 1650 nm ein Fehler von 0.66 m² m⁻². Die Genauigkeit, die mit dem physikalisch-basierten Ansatz erzielt wird, ist mit den Fehlerwerten empirischer Verfahren vergleichbar. Kalibrierte Reflexionsmodelle haben jedoch den Vorteil, dass sie generell auf unterschiedliche Sensortypen und unter wechselnden Rahmenbedingungen anwendbar sind.

Spekroradiometer (Field-Spec-II) gemessen und anschließend auf ihre Stoffkonzentrationen von Chlorophyll, Stickstoff, Wasser hin analysiert worden. Um die Übertragbarkeit auf Bilddaten abschätzen zu können, wird der Einfluss von spektraler Bandbreite und Sensorrauschen auf die Genauigkeit der statistischen Modelle untersucht. Die erzielten Ergebnisse zeigen, dass die Genauigkeit bei der Bestimmung der Blattinhaltsstoffe aus Laborspektren nur in geringem Maße durch eine verminderte spektrale Auflösung oder durch Sensorrauschen reduziert wird. Zum Hochkalieren von Blatt- auf Bestandsebene werden direkte Beziehung zwischen Blattanalysen und Bestandspektren erstellt. Die Chlorophyllkonzentration wird von allen Verfahren, jedoch mit unterschiedlicher, aber guter Genauigkeit bestimmt (kreuzvalidierter relater rmse 2.8-6.9 % des Mittelwerts; kreuzvalidiertes $R^2 = 0.90-0.46$). Die Bestimmung der Stickstoffkonzentration kann nur von drei Verfahren (wasserbereinigte Spektren, kontinuum-bereinigte Spektren und Ableitungsspekren) mit mäßiger Genauigkeit (rmse = 4.6-4.9 % des Mittelwerts; $R^2 = 0.52-0.57$) geleistet werden. Für den Blattwassergehalt können hingegen keine signifikanten Beziehungen gefunden werden. Selektierte Wellenlängen zur Schätzung von Chlorophyll befinden sich typischerweise im „red edge“ und im Randbereich der Chlorophyllabsorptionsbanden. Regressionsmodelle zur Schätzung von Stickstoff haben meist einen deutlichen Bezug zu den Absorptionsbanden von Chlorophyll oder Protein oder zum „red edge“. Die Verwendung der genannten Absorptionsbanden wird auch durch vergleichbare Untersuchungen anderer Vegetationstypen bestätigt. Die räumliche Verteilung der Blattchlorophyllkonzentration im Untersuchungsgebiet spiegelt die Nährstoffversorgung der Böden wider. So werden über basenreichen Basaltverwitterungslehmen im Mittel deutlich höhere Chlorophyllkonzentrationen beobachtet als auf den basenarmen Sanden des Mittleren Buntsandsteins.

### Appendix A: Field data

#### A1: Plots at Idarwald test site

<table>
<thead>
<tr>
<th>Plot-ID</th>
<th>Stand number</th>
<th>E</th>
<th>N</th>
<th>Stand age [yrs]</th>
<th>PBH [m]</th>
<th>SD [ha]</th>
<th>CO</th>
<th>LAI</th>
<th>H [m]</th>
<th>CD [m]</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>171a</td>
<td>2586713</td>
<td>5524702</td>
<td>46</td>
<td>0.92</td>
<td>622</td>
<td>0.45</td>
<td>3.81</td>
<td>18.0</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>59</td>
<td>155a</td>
<td>2587195</td>
<td>5524101</td>
<td>20</td>
<td>0.48</td>
<td>1489</td>
<td>0.65</td>
<td>5.47</td>
<td>9.3</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>65</td>
<td>165</td>
<td>2586664</td>
<td>5523910</td>
<td>68</td>
<td>1.15</td>
<td>511</td>
<td>0.40</td>
<td>3.31</td>
<td>29.3</td>
<td>4.9</td>
<td>ok</td>
</tr>
<tr>
<td>81</td>
<td>148b</td>
<td>2587394</td>
<td>5523645</td>
<td>18</td>
<td>0.49</td>
<td>1333</td>
<td>0.60</td>
<td>4.94</td>
<td>6.2</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>82</td>
<td>157a</td>
<td>2586433</td>
<td>5523369</td>
<td>135</td>
<td>1.57</td>
<td>278</td>
<td>0.50</td>
<td>2.48</td>
<td>30.6</td>
<td>7.2</td>
<td>ok</td>
</tr>
<tr>
<td>83</td>
<td>156a</td>
<td>2587024</td>
<td>5523430</td>
<td>10</td>
<td>0.27</td>
<td>2444</td>
<td>0.70</td>
<td>4.23</td>
<td>5.0</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>85</td>
<td>149a</td>
<td>2587282</td>
<td>5523455</td>
<td>49</td>
<td>0.81</td>
<td>800</td>
<td>0.50</td>
<td>2.63</td>
<td>23.0</td>
<td>-</td>
<td>recent thinning</td>
</tr>
<tr>
<td>100</td>
<td>158a</td>
<td>2586294</td>
<td>5523166</td>
<td>57</td>
<td>0.91</td>
<td>467</td>
<td>0.35</td>
<td>2.88</td>
<td>23.8</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>110</td>
<td>150b1</td>
<td>2586887</td>
<td>5522995</td>
<td>119</td>
<td>1.24</td>
<td>478</td>
<td>0.35</td>
<td>2.97</td>
<td>30.1</td>
<td>4.1</td>
<td>ok</td>
</tr>
<tr>
<td>111</td>
<td>144a</td>
<td>2587696</td>
<td>5523062</td>
<td>95</td>
<td>1.55</td>
<td>244</td>
<td>0.30</td>
<td>1.66</td>
<td>26.9</td>
<td>6.5</td>
<td>ok</td>
</tr>
<tr>
<td>115</td>
<td>158a</td>
<td>2586477</td>
<td>5522942</td>
<td>148</td>
<td>1.62</td>
<td>511</td>
<td>0.60</td>
<td>2.90</td>
<td>39.1</td>
<td>7.0</td>
<td>ok</td>
</tr>
<tr>
<td>118</td>
<td>125a1</td>
<td>2585601</td>
<td>5522574</td>
<td>48</td>
<td>1.36</td>
<td>267</td>
<td>0.40</td>
<td>2.66</td>
<td>33.7</td>
<td>6.8</td>
<td>ok</td>
</tr>
<tr>
<td>125</td>
<td>145</td>
<td>2587196</td>
<td>5522843</td>
<td>104</td>
<td>1.27</td>
<td>367</td>
<td>0.45</td>
<td>3.31</td>
<td>26.9</td>
<td>6.3</td>
<td>ok</td>
</tr>
<tr>
<td>129</td>
<td>150a2</td>
<td>2586913</td>
<td>5522671</td>
<td>119</td>
<td>1.38</td>
<td>378</td>
<td>0.40</td>
<td>2.24</td>
<td>27.3</td>
<td>6.1</td>
<td>ok</td>
</tr>
<tr>
<td>137</td>
<td>142a</td>
<td>2587548</td>
<td>5522577</td>
<td>34</td>
<td>0.81</td>
<td>955</td>
<td>0.55</td>
<td>4.08</td>
<td>13.0</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>138</td>
<td>126a</td>
<td>2585435</td>
<td>5522317</td>
<td>62</td>
<td>0.90</td>
<td>611</td>
<td>0.35</td>
<td>2.98</td>
<td>27.1</td>
<td>3.8</td>
<td>ok</td>
</tr>
<tr>
<td>144</td>
<td>111a</td>
<td>2586099</td>
<td>5522455</td>
<td>108</td>
<td>1.33</td>
<td>367</td>
<td>0.35</td>
<td>2.12</td>
<td>30.8</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>151</td>
<td>112a</td>
<td>2585852</td>
<td>5522190</td>
<td>111</td>
<td>1.09</td>
<td>578</td>
<td>0.60</td>
<td>2.42</td>
<td>26.6</td>
<td>5.6</td>
<td>ok</td>
</tr>
<tr>
<td>157</td>
<td>146a</td>
<td>2587324</td>
<td>5522264</td>
<td>70</td>
<td>1.15</td>
<td>600</td>
<td>0.55</td>
<td>5.07</td>
<td>26.3</td>
<td>4.5</td>
<td>ok</td>
</tr>
<tr>
<td>169</td>
<td>104</td>
<td>2586216</td>
<td>5521758</td>
<td>81</td>
<td>1.17</td>
<td>367</td>
<td>0.45</td>
<td>2.60</td>
<td>35.4</td>
<td>4.6</td>
<td>ok</td>
</tr>
<tr>
<td>187</td>
<td>125a2</td>
<td>2585752</td>
<td>5522695</td>
<td>64</td>
<td>1.06</td>
<td>422</td>
<td>0.30</td>
<td>2.52</td>
<td>26.0</td>
<td>4.5</td>
<td>ok</td>
</tr>
<tr>
<td>191</td>
<td>95a</td>
<td>2586627</td>
<td>5522094</td>
<td>80</td>
<td>1.63</td>
<td>411</td>
<td>0.55</td>
<td>2.70</td>
<td>30.3</td>
<td>6.3</td>
<td>ok</td>
</tr>
<tr>
<td>195</td>
<td>149a3</td>
<td>2587387</td>
<td>5523456</td>
<td>93</td>
<td>1.29</td>
<td>356</td>
<td>0.35</td>
<td>2.36</td>
<td>26.3</td>
<td>4.8</td>
<td>ok</td>
</tr>
<tr>
<td>208</td>
<td>162a</td>
<td>2587563</td>
<td>552251</td>
<td>69</td>
<td>1.09</td>
<td>456</td>
<td>0.40</td>
<td>2.53</td>
<td>31.9</td>
<td>5.0</td>
<td>ok</td>
</tr>
<tr>
<td>218</td>
<td>113b</td>
<td>2585104</td>
<td>5521825</td>
<td>53</td>
<td>1.07</td>
<td>433</td>
<td>0.50</td>
<td>3.27</td>
<td>21.4</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>219</td>
<td>128b</td>
<td>2584541</td>
<td>5521843</td>
<td>68</td>
<td>1.12</td>
<td>589</td>
<td>0.55</td>
<td>1.86</td>
<td>26.7</td>
<td>6.1</td>
<td>ok</td>
</tr>
</tbody>
</table>
## A1: Plots at Idarwald test site (continued)

<table>
<thead>
<tr>
<th>Plot-ID</th>
<th>Stand number</th>
<th>E</th>
<th>N</th>
<th>Stand age [yrs]</th>
<th>PBH [m]</th>
<th>SD [ha]</th>
<th>CO</th>
<th>LAI</th>
<th>H [m]</th>
<th>CD [m]</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td>118a2</td>
<td>2583737</td>
<td>5520076</td>
<td>87</td>
<td>0.91</td>
<td>722</td>
<td>0.35</td>
<td>2.71</td>
<td>25.0</td>
<td>3.9</td>
<td>ok</td>
</tr>
<tr>
<td>223</td>
<td>101c</td>
<td>2584336</td>
<td>5519335</td>
<td>106</td>
<td>1.21</td>
<td>433</td>
<td>0.35</td>
<td>2.28</td>
<td>27.5</td>
<td>3.9</td>
<td>outside HyMap stripe</td>
</tr>
<tr>
<td>252</td>
<td>106a</td>
<td>2585225</td>
<td>5520983</td>
<td>68</td>
<td>1.07</td>
<td>522</td>
<td>0.55</td>
<td>3.20</td>
<td>25.9</td>
<td>5.5</td>
<td>ok</td>
</tr>
<tr>
<td>254</td>
<td>115a</td>
<td>2584710</td>
<td>5520838</td>
<td>70</td>
<td>1.09</td>
<td>567</td>
<td>0.55</td>
<td>2.28</td>
<td>26.0</td>
<td>5.7</td>
<td>ok</td>
</tr>
<tr>
<td>263</td>
<td>116b</td>
<td>2584361</td>
<td>5520543</td>
<td>10</td>
<td>0.40</td>
<td>2044</td>
<td>0.70</td>
<td>5.22</td>
<td>6.5</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>271</td>
<td>107c</td>
<td>2585159</td>
<td>5520472</td>
<td>90</td>
<td>1.27</td>
<td>389</td>
<td>0.40</td>
<td>2.66</td>
<td>28.2</td>
<td>5.5</td>
<td>ok</td>
</tr>
<tr>
<td>283</td>
<td>69b</td>
<td>2583502</td>
<td>5519911</td>
<td>47</td>
<td>0.98</td>
<td>533</td>
<td>0.55</td>
<td>3.66</td>
<td>23.0</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>292</td>
<td>58a</td>
<td>2583839</td>
<td>5519562</td>
<td>64</td>
<td>1.19</td>
<td>411</td>
<td>0.40</td>
<td>3.35</td>
<td>25.1</td>
<td>5.1</td>
<td>ok</td>
</tr>
<tr>
<td>298</td>
<td>55a2</td>
<td>2583833</td>
<td>5519305</td>
<td>99</td>
<td>1.19</td>
<td>444</td>
<td>0.50</td>
<td>4.01</td>
<td>30.4</td>
<td>5.1</td>
<td>ok</td>
</tr>
<tr>
<td>301</td>
<td>55a3</td>
<td>2583717</td>
<td>5519228</td>
<td>120</td>
<td>1.28</td>
<td>344</td>
<td>0.40</td>
<td>3.06</td>
<td>35.1</td>
<td>6.1</td>
<td>ok</td>
</tr>
<tr>
<td>308</td>
<td>71a</td>
<td>2582893</td>
<td>5519122</td>
<td>20</td>
<td>0.57</td>
<td>1378</td>
<td>0.70</td>
<td>5.16</td>
<td>14.0</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>335</td>
<td>73a</td>
<td>2581995</td>
<td>5518293</td>
<td>44</td>
<td>0.82</td>
<td>955</td>
<td>0.65</td>
<td>4.67</td>
<td>22.2</td>
<td>-</td>
<td>ok</td>
</tr>
<tr>
<td>345</td>
<td>73c</td>
<td>2581881</td>
<td>5517925</td>
<td>126</td>
<td>1.49</td>
<td>278</td>
<td>0.40</td>
<td>2.23</td>
<td>36.6</td>
<td>6.3</td>
<td>ok</td>
</tr>
<tr>
<td>361</td>
<td>63a2</td>
<td>2581768</td>
<td>5517253</td>
<td>68</td>
<td>1.10</td>
<td>611</td>
<td>0.50</td>
<td>4.11</td>
<td>25.0</td>
<td>4.5</td>
<td>ok</td>
</tr>
<tr>
<td>389</td>
<td>108b</td>
<td>2584646</td>
<td>5520290</td>
<td>91</td>
<td>1.21</td>
<td>411</td>
<td>0.40</td>
<td>2.76</td>
<td>28.0</td>
<td>6.2</td>
<td>ok</td>
</tr>
<tr>
<td>391</td>
<td>110as</td>
<td>2584138</td>
<td>5519371</td>
<td>100</td>
<td>1.30</td>
<td>378</td>
<td>0.45</td>
<td>3.52</td>
<td>28.6</td>
<td>5.8</td>
<td>ok</td>
</tr>
</tbody>
</table>
### A2: Plots at Gerolstein test site

<table>
<thead>
<tr>
<th>Plot-ID</th>
<th>E</th>
<th>N</th>
<th>Owner</th>
<th>Stand number</th>
<th>Stand age [yrs]</th>
<th>Type of substrate</th>
<th>Base saturation</th>
<th>C_{\text{soil}} [mg/g dry matter]</th>
<th>C_{\text{w}} [% dry matter]</th>
<th>C_{\text{n}} [% dry matter]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS2</td>
<td>2548432</td>
<td>5568600</td>
<td>Rockeskyll</td>
<td>4d</td>
<td>51</td>
<td>BVL</td>
<td>high</td>
<td>3.32</td>
<td>132.8</td>
<td>1.16</td>
</tr>
<tr>
<td>GS3</td>
<td>2549121</td>
<td>5568205</td>
<td>Rockeskyll</td>
<td>4b</td>
<td>44</td>
<td>DLB</td>
<td>high</td>
<td>3.46</td>
<td>133.2</td>
<td>1.25</td>
</tr>
<tr>
<td>GS4</td>
<td>2548350</td>
<td>5568119</td>
<td>Rockeskyll</td>
<td>4e</td>
<td>55</td>
<td>Kol+</td>
<td>high</td>
<td>3.40</td>
<td>127.3</td>
<td>1.13</td>
</tr>
<tr>
<td>GS5</td>
<td>2550398</td>
<td>5566541</td>
<td>Pelm</td>
<td>20b</td>
<td>48</td>
<td>BVL</td>
<td>high</td>
<td>3.62</td>
<td>137.1</td>
<td>1.31</td>
</tr>
<tr>
<td>GS6</td>
<td>2548329</td>
<td>5566132</td>
<td>Pelm</td>
<td>23a</td>
<td>132</td>
<td>KVL</td>
<td>very high</td>
<td>2.99</td>
<td>139.5</td>
<td>1.18</td>
</tr>
<tr>
<td>GS7b</td>
<td>2548684</td>
<td>5565191</td>
<td>Pelm</td>
<td>19a</td>
<td>135</td>
<td>KVL</td>
<td>very high</td>
<td>3.23</td>
<td>139.5</td>
<td>1.29</td>
</tr>
<tr>
<td>GS8</td>
<td>2549466</td>
<td>5562358</td>
<td>State</td>
<td>307a</td>
<td>99</td>
<td>Vldu</td>
<td>low-medium</td>
<td>3.20</td>
<td>148.2</td>
<td>1.23</td>
</tr>
<tr>
<td>GS9</td>
<td>2550294</td>
<td>5563632</td>
<td>Pelm</td>
<td>9b</td>
<td>52</td>
<td>wDLT</td>
<td>medium</td>
<td>2.79</td>
<td>140.7</td>
<td>1.08</td>
</tr>
<tr>
<td>GS10</td>
<td>2550071</td>
<td>5563813</td>
<td>Pelm</td>
<td>11a</td>
<td>70</td>
<td>wDLT</td>
<td>medium</td>
<td>3.56</td>
<td>144.1</td>
<td>1.26</td>
</tr>
<tr>
<td>GS11</td>
<td>2548682</td>
<td>5564080</td>
<td>Pelm</td>
<td>17d</td>
<td>46</td>
<td>Ssm</td>
<td>very low</td>
<td>2.71</td>
<td>122.5</td>
<td>1.08</td>
</tr>
<tr>
<td>GS12</td>
<td>2549041</td>
<td>5564175</td>
<td>Pelm</td>
<td>16b</td>
<td>105</td>
<td>Ssm</td>
<td>very low</td>
<td>2.82</td>
<td>125.9</td>
<td>1.08</td>
</tr>
<tr>
<td>GS13</td>
<td>2549699</td>
<td>5563585</td>
<td>Pelm</td>
<td>9c</td>
<td>48</td>
<td>DLsm</td>
<td>low</td>
<td>3.20</td>
<td>149.0</td>
<td>1.25</td>
</tr>
<tr>
<td>GS14</td>
<td>2547944</td>
<td>5564210</td>
<td>Private</td>
<td>Fam. Moss</td>
<td>50</td>
<td>DLsm</td>
<td>low</td>
<td>3.42</td>
<td>134.7</td>
<td>1.22</td>
</tr>
</tbody>
</table>
Appendix B: Chemical analysis data

### B1: Needle samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>C&lt;sub&gt;A&lt;/sub&gt; [mg/g dry matter]</th>
<th>C&lt;sub&gt;B&lt;/sub&gt; [mg/g dry matter]</th>
<th>C&lt;sub&gt;N&lt;/sub&gt; [mg/g dry matter]</th>
<th>C&lt;sub&gt;W&lt;/sub&gt; [% dry matter]</th>
<th>C&lt;sub&gt;N&lt;/sub&gt; [% dry matter]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS2-1.1</td>
<td>3.09</td>
<td>1.73</td>
<td>4.82</td>
<td>145.6</td>
<td>1.20</td>
</tr>
<tr>
<td>GS2-1.3</td>
<td>3.74</td>
<td>2.15</td>
<td>5.89</td>
<td>120.7</td>
<td>1.22</td>
</tr>
<tr>
<td>GS2-2.1</td>
<td>2.31</td>
<td>1.46</td>
<td>3.77</td>
<td>145.8</td>
<td>1.15</td>
</tr>
<tr>
<td>GS2-2.3</td>
<td>3.19</td>
<td>2.03</td>
<td>5.22</td>
<td>109.6</td>
<td>1.15</td>
</tr>
<tr>
<td>GS2-3.1</td>
<td>2.72</td>
<td>1.68</td>
<td>4.40</td>
<td>156.5</td>
<td>1.18</td>
</tr>
<tr>
<td>GS2-3.3</td>
<td>3.25</td>
<td>1.96</td>
<td>5.21</td>
<td>118.4</td>
<td>1.06</td>
</tr>
<tr>
<td>GS3-1.1</td>
<td>2.23</td>
<td>0.58</td>
<td>2.81</td>
<td>147.5</td>
<td>1.26</td>
</tr>
<tr>
<td>GS3-1.3</td>
<td>2.63</td>
<td>0.82</td>
<td>3.46</td>
<td>120.4</td>
<td>1.21</td>
</tr>
<tr>
<td>GS3-2.1</td>
<td>2.51</td>
<td>0.61</td>
<td>3.12</td>
<td>161.6</td>
<td>1.29</td>
</tr>
<tr>
<td>GS3-2.3</td>
<td>0.96</td>
<td>1.52</td>
<td>2.49</td>
<td>122.5</td>
<td>1.22</td>
</tr>
<tr>
<td>GS3-3.1</td>
<td>2.43</td>
<td>0.63</td>
<td>3.06</td>
<td>138.4</td>
<td>1.31</td>
</tr>
<tr>
<td>GS3-3.3</td>
<td>2.38</td>
<td>0.68</td>
<td>3.06</td>
<td>108.7</td>
<td>1.19</td>
</tr>
<tr>
<td>GS4-1.1</td>
<td>2.25</td>
<td>0.57</td>
<td>2.82</td>
<td>143.1</td>
<td>1.22</td>
</tr>
<tr>
<td>GS4-1.3</td>
<td>3.41</td>
<td>1.30</td>
<td>4.71</td>
<td>108.5</td>
<td>1.11</td>
</tr>
<tr>
<td>GS4-2.1</td>
<td>2.33</td>
<td>1.32</td>
<td>3.65</td>
<td>143.2</td>
<td>1.01</td>
</tr>
<tr>
<td>GS4-2.3</td>
<td>3.25</td>
<td>1.67</td>
<td>5.12</td>
<td>112.7</td>
<td>1.06</td>
</tr>
<tr>
<td>GS4-3.1</td>
<td>2.45</td>
<td>1.07</td>
<td>3.51</td>
<td>138.6</td>
<td>1.25</td>
</tr>
<tr>
<td>GS4-3.3</td>
<td>3.20</td>
<td>1.65</td>
<td>4.85</td>
<td>117.8</td>
<td>1.13</td>
</tr>
<tr>
<td>GS8-1.1</td>
<td>2.16</td>
<td>0.52</td>
<td>2.68</td>
<td>159.1</td>
<td>1.34</td>
</tr>
<tr>
<td>GS8-1.3</td>
<td>3.32</td>
<td>0.97</td>
<td>4.29</td>
<td>131.5</td>
<td>1.27</td>
</tr>
<tr>
<td>GS8-2.1</td>
<td>2.59</td>
<td>0.69</td>
<td>3.27</td>
<td>155.0</td>
<td>1.40</td>
</tr>
<tr>
<td>GS8-2.3</td>
<td>2.43</td>
<td>0.72</td>
<td>3.15</td>
<td>114.6</td>
<td>1.23</td>
</tr>
</tbody>
</table>
## B1: Needle samples (continued)

<table>
<thead>
<tr>
<th>Sample</th>
<th>( C_A ) [mg/g dry matter]</th>
<th>( C_B ) [mg/g dry matter]</th>
<th>( C_{AB} ) [mg/g dry matter]</th>
<th>( C_W ) [% dry matter]</th>
<th>( C_N ) [% dry matter]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS8-3.1</td>
<td>2.37</td>
<td>0.59</td>
<td>2.97</td>
<td>147.4</td>
<td>1.46</td>
</tr>
<tr>
<td>GS8-3.3</td>
<td>2.68</td>
<td>0.79</td>
<td>3.47</td>
<td>114.9</td>
<td>1.16</td>
</tr>
<tr>
<td>GS9-1.1</td>
<td>1.96</td>
<td>0.47</td>
<td>2.43</td>
<td>158.4</td>
<td>1.27</td>
</tr>
<tr>
<td>GS9-1.3</td>
<td>2.38</td>
<td>0.62</td>
<td>2.99</td>
<td>127.8</td>
<td>1.24</td>
</tr>
<tr>
<td>GS9-2.1</td>
<td>1.57</td>
<td>0.39</td>
<td>1.97</td>
<td>151.8</td>
<td>1.00</td>
</tr>
<tr>
<td>GS9-2.3</td>
<td>2.11</td>
<td>0.61</td>
<td>2.72</td>
<td>127.5</td>
<td>1.05</td>
</tr>
<tr>
<td>GS9-3.1</td>
<td>1.90</td>
<td>0.48</td>
<td>2.39</td>
<td>153.0</td>
<td>1.19</td>
</tr>
<tr>
<td>GS9-3.3</td>
<td>2.80</td>
<td>0.86</td>
<td>3.66</td>
<td>118.3</td>
<td>1.30</td>
</tr>
<tr>
<td>GS10b-1.1</td>
<td>1.94</td>
<td>0.51</td>
<td>2.44</td>
<td>146.2</td>
<td>1.31</td>
</tr>
<tr>
<td>GS10b-1.3</td>
<td>2.94</td>
<td>0.96</td>
<td>3.90</td>
<td>113.3</td>
<td>1.34</td>
</tr>
<tr>
<td>GS10b-2.1</td>
<td>1.85</td>
<td>0.46</td>
<td>2.31</td>
<td>156.9</td>
<td>1.33</td>
</tr>
<tr>
<td>GS10b-2.3</td>
<td>2.54</td>
<td>0.76</td>
<td>3.30</td>
<td>129.6</td>
<td>NaN</td>
</tr>
<tr>
<td>GS10b-3.1</td>
<td>1.73</td>
<td>0.42</td>
<td>2.16</td>
<td>166.3</td>
<td>1.26</td>
</tr>
<tr>
<td>GS10b-3.3</td>
<td>2.57</td>
<td>0.74</td>
<td>3.31</td>
<td>124.6</td>
<td>1.22</td>
</tr>
<tr>
<td>GS16-1.1</td>
<td>1.92</td>
<td>0.60</td>
<td>2.52</td>
<td>161.3</td>
<td>1.23</td>
</tr>
<tr>
<td>GS16-1.3</td>
<td>2.26</td>
<td>0.82</td>
<td>3.08</td>
<td>128.2</td>
<td>1.09</td>
</tr>
<tr>
<td>GS16-2.1</td>
<td>1.92</td>
<td>0.51</td>
<td>2.43</td>
<td>153.7</td>
<td>1.19</td>
</tr>
<tr>
<td>GS16-2.3</td>
<td>2.90</td>
<td>1.27</td>
<td>4.17</td>
<td>138.8</td>
<td>1.16</td>
</tr>
<tr>
<td>GS16-3.1</td>
<td>2.19</td>
<td>0.58</td>
<td>2.76</td>
<td>166.7</td>
<td>1.31</td>
</tr>
<tr>
<td>GS16-3.3</td>
<td>2.83</td>
<td>0.91</td>
<td>3.74</td>
<td>140.4</td>
<td>1.37</td>
</tr>
<tr>
<td>GS17-1.1</td>
<td>1.64</td>
<td>0.39</td>
<td>2.03</td>
<td>148.8</td>
<td>1.16</td>
</tr>
<tr>
<td>GS17-1.3</td>
<td>2.24</td>
<td>0.64</td>
<td>2.88</td>
<td>123.5</td>
<td>1.10</td>
</tr>
<tr>
<td>GS17-2.1</td>
<td>1.31</td>
<td>0.40</td>
<td>1.71</td>
<td>136.1</td>
<td>0.91</td>
</tr>
<tr>
<td>GS17-2.3</td>
<td>1.58</td>
<td>0.55</td>
<td>2.13</td>
<td>133.5</td>
<td>0.88</td>
</tr>
<tr>
<td>GS17-3.1</td>
<td>2.70</td>
<td>0.71</td>
<td>3.41</td>
<td>164.2</td>
<td>1.30</td>
</tr>
<tr>
<td>GS17-3.3</td>
<td>2.61</td>
<td>0.90</td>
<td>3.51</td>
<td>138.2</td>
<td>1.11</td>
</tr>
</tbody>
</table>
## B1: Needle samples (continued)

<table>
<thead>
<tr>
<th>Sample</th>
<th>$C_A$ [mg/g dry matter]</th>
<th>$C_B$ [mg/g dry matter]</th>
<th>$C_{AB}$ [mg/g dry matter]</th>
<th>$C_W$ [% dry matter]</th>
<th>$C_N$ [% dry matter]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS18-1.1</td>
<td>2.10</td>
<td>0.54</td>
<td>2.64</td>
<td>155.0</td>
<td>1.18</td>
</tr>
<tr>
<td>GS18-1.3</td>
<td>2.64</td>
<td>0.81</td>
<td>3.45</td>
<td>128.7</td>
<td>1.13</td>
</tr>
<tr>
<td>GS18-2.1</td>
<td>2.28</td>
<td>0.62</td>
<td>2.91</td>
<td>169.2</td>
<td>1.40</td>
</tr>
<tr>
<td>GS18-2.3</td>
<td>3.11</td>
<td>0.95</td>
<td>4.06</td>
<td>127.5</td>
<td>1.28</td>
</tr>
<tr>
<td>GS18-3.1</td>
<td>2.50</td>
<td>0.70</td>
<td>3.20</td>
<td>146.2</td>
<td>1.32</td>
</tr>
<tr>
<td>GS18-3.3</td>
<td>2.83</td>
<td>0.85</td>
<td>3.68</td>
<td>138.0</td>
<td>1.24</td>
</tr>
<tr>
<td>GS19-1.1</td>
<td>1.62</td>
<td>0.40</td>
<td>2.02</td>
<td>126.2</td>
<td>0.98</td>
</tr>
<tr>
<td>GS19-1.3</td>
<td>2.06</td>
<td>0.61</td>
<td>2.66</td>
<td>113.7</td>
<td>1.07</td>
</tr>
<tr>
<td>GS19-2.1</td>
<td>1.99</td>
<td>0.51</td>
<td>2.50</td>
<td>141.2</td>
<td>1.16</td>
</tr>
<tr>
<td>GS19-2.3</td>
<td>2.13</td>
<td>0.61</td>
<td>2.74</td>
<td>112.7</td>
<td>1.06</td>
</tr>
<tr>
<td>GS19-3.1</td>
<td>1.75</td>
<td>0.45</td>
<td>2.19</td>
<td>130.6</td>
<td>1.12</td>
</tr>
<tr>
<td>GS19-3.3</td>
<td>2.22</td>
<td>0.66</td>
<td>2.88</td>
<td>110.4</td>
<td>1.06</td>
</tr>
<tr>
<td>GS21-1.1</td>
<td>2.16</td>
<td>0.54</td>
<td>2.70</td>
<td>137.4</td>
<td>1.17</td>
</tr>
<tr>
<td>GS21-1.3</td>
<td>2.70</td>
<td>0.80</td>
<td>3.50</td>
<td>112.5</td>
<td>1.09</td>
</tr>
<tr>
<td>GS21-2.1</td>
<td>1.76</td>
<td>0.45</td>
<td>2.21</td>
<td>142.4</td>
<td>1.09</td>
</tr>
<tr>
<td>GS21-2.3</td>
<td>2.11</td>
<td>0.64</td>
<td>2.74</td>
<td>107.9</td>
<td>1.06</td>
</tr>
<tr>
<td>GS21-3.1</td>
<td>1.66</td>
<td>0.48</td>
<td>2.14</td>
<td>138.7</td>
<td>1.07</td>
</tr>
<tr>
<td>GS21-3.3</td>
<td>1.76</td>
<td>0.53</td>
<td>2.29</td>
<td>116.6</td>
<td>0.99</td>
</tr>
<tr>
<td>GS24-1.1</td>
<td>1.90</td>
<td>0.45</td>
<td>2.36</td>
<td>163.2</td>
<td>1.29</td>
</tr>
<tr>
<td>GS24-1.3</td>
<td>2.39</td>
<td>0.74</td>
<td>3.13</td>
<td>129.0</td>
<td>1.19</td>
</tr>
<tr>
<td>GS24-2.1</td>
<td>2.35</td>
<td>0.62</td>
<td>2.97</td>
<td>150.4</td>
<td>1.32</td>
</tr>
<tr>
<td>GS24-2.3</td>
<td>2.46</td>
<td>0.69</td>
<td>3.15</td>
<td>141.8</td>
<td>1.22</td>
</tr>
<tr>
<td>GS24-3.1</td>
<td>2.20</td>
<td>0.57</td>
<td>2.76</td>
<td>170.8</td>
<td>1.28</td>
</tr>
<tr>
<td>GS24-3.3</td>
<td>2.51</td>
<td>0.74</td>
<td>3.24</td>
<td>138.6</td>
<td>1.18</td>
</tr>
</tbody>
</table>
Appendix C: Matlab code

C1: INFORM model

function [r_forest,r_soil,r_understorey,r_c_inf,co,C,G,r_needle,t_needle,t_s,t_o]=inform_liberty(PARA_LIBERTY,PARA_INFORM,teta_o,teta_s,phi,skyl,r_soil,r_wood,data,liberty_data_hymap);

% INFORM_LIBERTY - The INvertible FOrest Reflectance Model coupled with LIBERTY

% INFORM (Atzberger, 2000; Schlerf & Atzberger, 2005) simulates the bi-directional reflectance
% of forest stands between 400 and 2500 nm. INFORM is essentially an innovative combination of
% FLIM (Rosema et al., 1992), SAIL (Verhoef, 1984), and PROSPECT (Jacquemoud & Baret, 1990) or
% LIBERTY (Dawson et al., 1998).

% In: Buchroithner (Ed.): A decade of trans-european remote sensing cooperation. Proceedings

% Schlerf, M. & Atzberger, C. (2005): Inversion of a forest reflectance model to estimate biophysical
% canopy variables from hyperspectral remote sensing data. Submitted to Remote Sensing of Environment.

% Rosema, A., Verhoef, W., Noorbergen, H. 1992: A new forest light interaction model in support of forest

% Environment, 34: 75-91.


% Verhoef, W. 1984: Light scattering by leaf layers with application to canopy reflectance modeling: The
C1: INFORM model (continued)

% Basic version of INFORM: Clement Atzberger, 1999
% Implementation of LIBERTY: Sebastian Mader, 2002
% INFORM modifications and validation: Martin Schlerf, 2004

% _____________________________________________________________________________________________________________
% SUBROUTINES
% _____________________________________________________________________________________________________________

% M-File Function
% sail_background.m SAIL-PROSPECT-SOIL to compute background reflectance
% sail_inf.m SAIL-PROSPECT-SOIL to compute infinite crown reflectance
% sail_t_s.m SAIL-PROSPECT-SOIL to compute crown transmittance for sun direction
% sail_t_o.m SAIL-PROSPECT-SOIL to compute crown transmittance for observation direction
% liberty.m Needle reflectance model LIBERTY
% prospect.m Learf reflectance model PROSPECT (to compute understorey leaf reflectance)
% S13AAF.M Integral
% TAV.M Refraction index

% _____________________________________________________________________________________________________________

% INPUT VARIABLES

% PARA_LIBERTY: Needle Input Parameters

% Variable Designation Unit Default value
% Cell diameter d µm 40
% Intercellular air space i arbitrary 0.045
% Leaf thickness t arbitrary 1.6
% Baseline absorption b arbitrary 0.0006
% Albino absorption a arbitrary 2
% Chlorophyll content cab mg m-2 200
% Water content cw g m-2 100
% Lignin and cellulose content cl g m-2 40
% Nitrogen content cp g m-2 1
C1: INFORM model (continued)

% PARA_INFORM: Canopy Input Parameters

% Variable Designation  Unit  Default value
% Single tree leaf area index  lai  m2 m-2  7
% Leaf area index of understorey  laiu  m2 m-2  0.1
% PROSPECT parameter  vaiprosu  arbitrary  2
% Stem density  sd  ha-1  650
% Tree height  h  m  20
% Crown diameter  cd  m  4.5
% Average leaf angle of tree canopy  ala  deg  55
% Scale factor for soil reflectance  scale  fraction  1

% External Input Parameters

% Variable Designation  Unit  Default value
% Sun zenith angle  teta_s  deg  30
% Observation zenith angle  teta_o  deg  0
% Azimuth angle  phi  deg  0
% Fraction of diffuse radiation  skyl  fraction  0.1

% Other Input Data

% Variable Designation
% r_soil  Soil spectrum
% liberty_data_hymap  LIBERTY coefficients resampled to HyMap wavebands

% OUTPUT VARIABLES

% Variable Designation
% Forest reflectance  r_forest
% Soil reflectance  r_soil
% Understorey reflectance  r_understorey
% Infinite canopy reflectance  r_c_inf
% Canopy cover  co
% Crown factor  C
% Ground factor  G
C1: INFORM model (continued)

% Needle reflectance                   r_needle
% Needle transmittance                  t_needle
% Crown transmittance for teta_s        t_s
% Crown transmittance for teta_o        t_o
% _____________________________________________________________________________________________________________

lai=PARA_INFORM(1);
laiu=PARA_INFORM(2);
vaiproso=PARA_INFORM(4);
sd=PARA_INFORM(6);
h=PARA_INFORM(7);
cd=PARA_INFORM(8);
ala=PARA_INFORM(9);
scale=PARA_INFORM(10);
% _____________________________________________________________________________________________________________
% ...  % Computation of needle reflecance and transmittance for HyMap wavebands by LIBERTY
% [r_needle,t_needle]=liberty(PARA_LIBERTY', liberty_data_hymap);
% [rt]=prospect(1.2,40,0.015,data); r_needle=rt(:,1); t_needle=rt(:,2);
% Computation of infinitive crown reflectance for a very dense forest canopy with LAI=15 (hot=0, N=1.5)
r_c_inf=sail_inf(15,ala,0,teta_o,teta_s,phi,skyl,r_understorey,r_needle,t_needle);
% _____________________________________________________________________________________________________________
% Computation of spectral components (SAIL, PROSPECT, LIBERTY models)
% Scaling of soil spectrum (to account for effects due to shadow and soil moisture)
r_soil=scale*r_soil;
% Computation of understorey reflectance for dicotyledoneae (ala=45, hot=0, cab=30, cw=0.05, N=2.5)
r_r_understorey=sail_background(laiu,45,0,30,0.05,vaiproso,teta_o,teta_s,phi,skyl,r_soil,data);
% Computation of needle reflecance and transmittance for HyMap wavebands by LIBERTY
% [r_needle,t_needle]=liberty(PARA_LIBERTY', liberty_data_hymap);
% [rt]=prospect(1.2,40,0.015,data); r_needle=rt(:,1); t_needle=rt(:,2);

% Ground coverage (FLIM model)
% Computation of the average horizontal area of a single tree crown in hectare (k) corrected by the factor 'adapt'.
C1: INFORM model (continued)

% 'adapt' was adjusted in such a way, that modelled values of 'co' agreed with measured values of 'co' for a given
% stand structure (as defined by 'cd' and 'sd'). To compare measured with modelled stands for certain values of canopy
% lai (lai_c), an empirical relation between 'lai_c' and 'co' was derived from field data at Idarwald site
% (co=0.078*lai_c+0.23, n=40).

% adapt=1; % no correction
adapt=0.6;

k=adapt*(pi*(cd/2).^2)/10000;

% angles (degree) to angles (radian)
teta_o=teta_o*pi/180;
teta_s=teta_s*pi/180;
phi=phi*pi/180;

% Observed ground coverage by crowns (co) under observation zenith angle teta_o
co=1-exp(-k*sd/cos(teta_o));

% Ground coverage by shadow (cs) under a solar zenith angle teta_s
cs=1-exp(-k*sd/cos(teta_s));

% Geometrical factor (g) depending on the illumination and viewing geometry
g=((tan(teta_o)).^2+(tan(teta_s)).^2-2*tan(teta_o)*tan(teta_s)*cos(phi)).^(0.5);

% Correlation coefficient (p)
p=exp(-g*h/cd);

% Ground surface fractions (FLIM model)
% Tree crowns with shadowed background (Fcd)
Fcd=co*cs+p*(co*(1-co)*cs*(1-cs)).^(0.5);

% Tree crowns with sunlit background (Fcs)
Fcs=co*(1-cs)-p*(co*(1-co)*cs*(1-cs)).^(0.5)
% Shadowed open space (Fod)
Fod=(1-co)*cs-p*(co*(1-co)*cs*(1-cs)).^(0.5);

% Sunlit open space (Fos)
Fos=(1-co)*(1-cs)+p*(co*(1-co)*cs*(1-cs)).^(0.5);

% Crown transmittance (SAIL model)
% Angles (radian) to angles (degree)
teta_s=teta_s*180/pi;
teta_o=teta_o*180/pi;
% Crown transmittance in sun direction (t_s)
t_s=sail_t_s(lai,ala,0,teta_s,skyl,r_understorey,teta_o,phi,r_needle,t_needle);
% Crown transmittance in observation direction (t_o)
t_o=sail_t_o(lai,ala,0,teta_o,skyl,r_understorey,teta_s,phi,r_needle,t_needle);

% Forest reflectance (FLIM model)
% Ground factor (G); that is ground contribution to scene reflectance
G=Fod.*t_s.*t_o+Fcs.*t_o+Fod.*t_s+Fos;
% Crown factor (C); that is crown contribution to scene reflectance
C=(1-t_s.*t_o)*cs*co; % Original formula
% Forest reflectance
r_forest=r_c_inf.*C+r_understorey.*G;
C2: Inversion routine

% NN_INFORM
% Matlab Neural Network Toolbox 4, page 5-66 - 5-70
% 24.08.04
% Martin Schlerf

% Load calibration data
% PARA_INFORM: Simulated INFORM parameters
% R_FOREST: Simulated forest canopy reflectances
% lambda: HyMap wavebands
load data_cal

% Load validation parameters (LAI_val, CO_val, SD_val) and
% HyMap spectra (r)
load data_val

% Position of varying parameters
Par=[1 6 8];

% Define target
t=PARA_INFORM(:,Par)';

% Pre-standardise target
[tn,meant,stdt]=prestd(t);

% Divide target into training and test subsets
Q=length(tn);
iitr=[1:4:Q 3:4:Q];
iite=[2:4:Q 4:4:Q];
ttr=tn(:,iitr);
test.T=tn(:,iite);
% Define wavebands
wl=[837.2 1148.4];

for K=1:length(wl)
    pos(K)=find(lambda==wl(K));
end

% Define input vector
p=R_FOREST(pos,:);

% Pre-standardise input
[pn,meanp,stdp]=prestd(p);

% Divide input into training and test subsets
test.P=pn(:,iite);
ptr=pn(:,iitr);

% Network architecture: 2-layer network, tan-sigmoidal transfer
% function in the hidden layer, linear transfer function in the
% output layer, 1 neurons in the hidden layer, 3 output neurons
% (3 parameters to be estimated), resilient backpropagation algorithm
% (RPROP) for training.
net=newff(minmax(ptr),[1 3],{'tansig' 'purelin'},'trainrp');

% Definition of additional training specifications
net.trainParam.show=5;
net.trainParam.epochs=5000;

% Train network
[net,tr]=train(net,ptr,ttr,[],[],test,[]);

% Pre-standardise HyMap reflectances
[rn]=trastd(r(pos,:),meanp,stdp);
C2: Inversion routine (continued)

% Estimation of parameters
an = sim(net, rn);

% Unnormalise estimated parameters
[a] = poststd(an, meant, stdt);

% Compute CO (crown cover) from SD (stem density) and CD (crown diameter)
adapt = 0.6;
teta_o = 0; teta_o = teta_o * pi / 180;
teta_s = 30; teta_s = teta_s * pi / 180;
phi = 0;
CD_est = a(3, :);
k = adapt * (pi * (CD_est / 2).^2) / 10000;
SD_est = a(2, :);
CO_est = 1 - exp(-k * SD_est / cos(teta_o));
LAI_est = a(:, :) .* CO;

% Compute validation statistics
r2_LAI = corrcoef(LAI_est, LAI_val).^2; r2_2_LAI = r2_LAI(1, 2);
rmse_LAI = (mean((LAI_est - LAI_val).^2))^0.5;
r2_CO = corrcoef(CO_est, CO_val).^2; r2_2_CO = r2_CO(1, 2);
rmse_CO = (mean((CO_est - CO_val).^2))^0.5;
r2_SD = corrcoef(SD_est, SD_val).^2; r2_2_SD = r2_SD(1, 2);
rmse_SD = (mean((SD_est - SD_val).^2))^0.5;
I would like to acknowledge the following people for their highly appreciated support during the course of this research work:

My supervisor, Prof. Dr. Joachim Hill, for his confidence and advice during the past years. The work drew substantial benefit from his expertise; in particular, his experience in radiometric data processing was indispensable for the correction of the hyperspectral imagery.

My second supervisor, Prof. Dr. Willy Werner, for his interest in the topic of the work, for thorough discussions on ecological concerns and for providing the chemical analysis of leaf samples.

Assistant Prof. Dr. Clement Atzberger (now at ITC, Enschede) for sharing his knowledge of concepts and implementations, and for his enthusiasm, advice and encouragement during the past five years.

My colleagues at RSD for their manifold support and for the pleasant and motivating atmosphere, particularly:

Jun. Prof. Dr. Michael Vohland for his precursory studies at the Idarwald site, Dr. Thomas Udelhoven for his help with all kinds of statistical concerns and Dr. Thomas Jarmer for his assistance in spectrometric measurements.

All student staff who contributed to different aspects of the work. In particular Henning Buddenbaum who significantly supported the research, especially by providing programming codes with polished graphical interfaces; Sebastian Mader for his support in data processing and programming; Samuel Bärisch for his help with fieldwork at the Idarwald and Gerolstein test sites, for solving any kind of technical problem and for never losing orientation even in the deepest forest.

Stephan Seeling for the opportunities he made available to me in the WARELA project.

Hermine Marx for her outstanding talent to organize stuff and to make things work.

Clement Atzberger, Michael Vohland and Achim Röder for correcting the manuscript.

Achim Röder and Patrick Hostert for making life in Trier enjoyable.

Andreas Müller, Andrea Hausold and Stefanie Holzwarth (DLR, Oberpfaffenhofen) for organising flight campaigns.

Dr. habil. Gebhard Schüler (FAWF, Trippstadt), Dr. Werner Schwind and some brave woodmen for their support during fieldwork at Gerolstein test site.

The research of this work was partly carried out within the framework of the Sonderforschungsbereich 522 “Umwelt und Region” and the research project “Ableitung biochemischer Forstparameter mit Bildspektrometerdaten”, both funded by the Deutsche Forschungsgemeinschaft (DFG). Additionally, the following organisations contributed to this research work: Deutsche Luft- und Raumfahrtgesellschaft (DLR); Forschungsanstalt für Waldökologie und Forstwirtschaft Rheinland-Pfalz (FAWF); University of Trier: Geobotany Department, Department of Biogeography and Remote Sensing Department (RSD).

Finally, I would like to thank my partner Eve who gave me moral support during the whole time of the work.
Lebenslauf

Name: Martin Schlerf
Wohnort: Deutschherrenstrasse 48a, D-54290 Trier
Geburtsdatum: 23.06.1972
Geburtsort: Friedberg (Hess.)
Staatsangehörigkeit: Deutschland

Schulbildung
08/1982 – 07/1991 Ernst-Ludwig-Gymnasium, Bad Nauheim
08/1978 – 07/1982 Grundschule Wettertal, Bad Nauheim

Studium
04/2002 – 09/2003 Promotionsstudium, Universität Trier
Abschluss: Professional Master
Abschluss: Diplom-Geograph
Thema der Abschlussarbeit: GIS-gestützte Untersuchung von Permafrost und Hanginstabilitäten im Gornergratgebiet, Wallis

Berufstätigkeit
09/2003 – 2006 Wissenschaftlicher Mitarbeiter im EU-Projekt WARELA, Universität Trier
04/2003 – 2006 Wissenschaftlicher Mitarbeiter im Fach Fernerkundung, Universität Trier
10/2000 – 03/2003 Wissenschaftlicher Mitarbeiter im Sonderforschungsbereich 522 „Umwelt und Region“, Universität Trier