Prediction of Foliar Biochemistry in a Boreal Forest Canopy Using Imaging Spectroscopy and LiDAR Data

Kemal Gökkaya

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Forest Resources and Environmental Conservation

Valerie A. Thomas, Chair
Thomas L. Noland
Randolph H. Wynne
John R. Seiler

July 20, 2012
Blacksburg, Virginia

Keywords: imaging spectroscopy, LiDAR, macronutrients, N:P ratio, boreal forest

Prediction of Foliar Biochemistry in a Boreal Forest Canopy Using Imaging Spectroscopy and LiDAR Data

Kemal Gökkaya

ABSTRACT

The use of satellite and airborne remote sensing data to predict foliar macronutrients and pigments for a boreal mixedwood forest composed of black and white spruce, balsam fir, northern white cedar, white birch, and trembling aspen was investigated. Specifically, imaging spectroscopy (IS) and light detection and ranging (LiDAR) are used to model the foliar N:P ratio, macronutrients (N, P, K, Ca, Mg) and chlorophyll. Measurement of both foliar macronutrients and foliar chlorophyll provide critical information about plant physiological and nutritional status, stress, as well as ecosystem processes such as carbon (C) exchange (photosynthesis and net primary production), decomposition and nutrient cycling. Results show that airborne and spaceborne IS data explained approximately 70% of the variance in the canopy N:P ratio with predictions errors of less than 8% in two consecutive years. LiDAR models explained more than 50% of the variance in the canopy N:P ratio with similar predictions errors. Predictive models using spaceborne Hyperion IS data were developed with adjusted R^2 values of 0.73, 0.72, 0.62, 0.25, and 0.67 for N, P, K, Ca and Mg, respectively. The LiDAR model explained 80% of the variance in canopy Ca concentration with an RMSE of less than 10%, suggesting strong correlations between forest height and Ca. Two IS derivative indices emerged as good predictors of chlorophyll across time and space. When the models of these two indices with the same parameters as generated from Hyperion data were applied to other years' data for chlorophyll concentration prediction, they could explain 71, 63 and 6% and 61, 54 and 8 % of the variation in chlorophyll concentration in 2002, 2004 and 2008, respectively with prediction errors ranging from 11.7% to 14.6%. Results demonstrate that the N:P ratio, N, P, K, Mg and chlorophyll can be modeled by spaceborne IS data and Ca can only be predicted by LiDAR data in the canopy of this forest. The ability to model the N:P ratio and macronutrients using spaceborne Hyperion data demonstrates the potential for mapping them at the canopy scale across larger geographic areas and being able to integrate them in future studies of ecosystem processes.

Acknowledgments

First of all, I would like to thank my advisor, Valerie Thomas, for guiding me through my work. This dissertation would not have been possible without her help, advice and patience. Secondly, I thank my committee members. In particular, Dr. Thomas Noland was always available to provide data and suggestions, which improved the quality of the research presented here. Drs. Randolph Wynne and John Seiler had useful critiques, comments and suggestions.

Dr. Carolyn Copenhaver deserves special recognition as she was so generous to offer advice, assistance and share her personal experience with me to prepare me to be a better scientist. I would also like to thank Drs. Phil Radtke, Yang Shao and Charles Sabatia for helping me with statistical and technical issues. I thank Drs. Paul Treitz and Harry McCaughey for improving the quality of my manuscripts by providing feedback in a timely manner. I also would like to thank Dr. Nicholas Coops for being available to discuss science at conferences.

I cannot stress the importance of the support of fellow graduate students, namely Nilam Kayastha, Jess Walker, Alicia Peduzzi, Asim Banskota, Rupesh Shreshta, Valquiria Quirino, Susmita Sen, and Arvind Bhuta. Thank you all. Blacksburg, CEARS lab and the department was a more fun place because of you.

The wonderful staff of the department made it a pleasure to be a student there. Sue Snow is an exceptional person and every workplace needs someone like her. In addition, I need to thank Stacey Kuhar, Tracey Sherman, Kathie Hollandsworth for helping with travel arrangements, any kind of paperwork, website preparation and the small chats we've had in the halls with their smiles on their faces. The department chair, Dr. Janaki Alavalapati was always very open to students' ideas, very approachable and supportive. Thank you Dr. Alavalapati for epitomising an exemplary administrator and mentor.

At Virginia Tech, I had the opportunity to do things beyond my dissertation research, which prepared me to become an academician. I owe this firstly to my advisor, Dr. Valerie Thomas, for bringing it to my attention early on during my tenure at the university, and secondly and most importantly to Dr. Karen DePauw, the dean of the Graduate School. Thank you so much Dr. DePauw for enriching my graduate education, a wonderful trip to Switzerland and Italy, stressing

the importance of not just “looking” but also “seeing” and using the metaphor of observing the “doors and windows” to understand and appreciate the cultures we get to see, and being easily accessible and approachable whenever needed. College of Natural Resources and the Environment dean, Dr. Paul Winistorfer, also deserves special thanks for covering part of my travel expenses for the trip.

Lastly and most importantly I must thank my family. My mother, Ülkü, has supported and waited patiently for many years to see me graduate and become a doctor. She has always supported me in whatever I wanted to pursue in life. My dear wife, Sheila, was always there to help me and create the environment for me to be able to work. My beautiful son, Eren, rendered my doctorate study more enjoyable by his presence. My farmer uncle and maternal grandparents and the pastoral lifestyle I’ve had the chance to enjoy during my early childhood planted the seeds of love and respect towards nature, which eventually led to a formal study of it. Therefore, I owe my thanks to them as well. I cannot ignore the contribution of all the teachers and mentors through my elementary, middle, high school and undergraduate and master’s level education. Without your efforts, I would not have accomplished this task.

And of course, I would like to give my thanks to God, for giving me the strength, the health, the patience and the endurance to persevere during difficult times.

I thank you all wholeheartedly.

Table of Contents

ABSTRACT ii
Acknowledgments iii
Table of Contents v
List of Figures vii
List of Tables viii
Chapter 1 General Introduction and Objectives 1
1.1. Introduction 1
1.1.1. Importance of Foliar Biochemicals 1
1.1.2. The importance of the boreal forest for canopy biochemistry analysis 2
1.1.3. Remote Sensing of Foliar Biochemicals 4
1.1.3.1. Strategic Use of Remote Sensing in Ecology 4
1.1.3.2. Importance of Remotely Sensing Foliar Biochemicals at the Canopy Level 4
1.1.3.3. Imaging Spectroscopy and Light Detection and Ranging for Canopy Biochemistry Estimation 5
1.2. Objectives 10
1.3. Importance of research 11
1.4. References 12
Chapter 2 Canopy nitrogen to phosphorus ratio mapping at two different scales in a boreal mixedwood forest 19
Abstract 19
2.1. Introduction 20
2.2. Materials and Methods 24
2.2.1. Background 24
2.2.2. Study Site 25
2.2.3. Forest mensuration data 27
2.2.4. Biochemical sampling and analysis 27
2.2.5. Remote sensing data – preprocessing 28
2.2.6. Remote sensing data – generation of hyperspectral indices and LiDAR metrics 30
2.2.7. Statistical analyses for model development 31
2.2.8. Potential source of error 33
2.2.9. Assessment of model performance and consideration of the MAUP 33
2.3. Results 34
2.4. Discussion 39

2.4.1. Effect of temporal variation on prediction of the canopy N:P ratio 39
2.4.2. Comparison of the two methods of predicting the canopy N:P ratio..... 40
2.4.3. Effect of the MAUP on model performance..... 41
2.4.4. Predictive models..... 42
2.4.5. Spatial distribution of the canopy N:P ratio at GRFS and implications 44
2.5. Conclusions..... 48
2.6. Acknowledgements..... 49
2.7. References..... 50
Chapter 3 Prediction of macronutrients using spaceborne imaging spectroscopy and LiDAR data in a mixedwood boreal forest canopy 57
Abstract..... 57
3.1. Introduction..... 58
3.2. Materials and Methods..... 62
3.2.1. Study Site 62
3.2.2. Field data..... 63
3.2.3. Remote sensing data 65
3.2.4. Model development 65
3.2.5. Relationships between macronutrients and canopy structure 66
3.2.6. Regression Analysis..... 67
3.3. Results..... 68
3.4. Discussion..... 75
3.4.1. IS predictive models 75
3.4.2. Relationships between macronutrients and canopy structure and LiDAR models..... 78
3.5. Conclusions..... 79
3.6. Acknowledgements..... 80
3.7. References..... 81
Chapter 4 Testing the robustness of predictive models for chlorophyll generated from spaceborne imaging spectroscopy data in mixedwood boreal forest canopy 86
Abstract..... 86
4.1. Introduction..... 87
4.2. Materials and Methods..... 89
4.2.1. Study sites 89
4.2.2. Biochemical sampling and analysis 90
4.2.3. Remote sensing data 92
4.2.4. Generation of hyperspectral indices..... 93

4.2.5. Statistical analyses, model development and testing robustness of models and indices 95
4.3. Results ... 97
4.4. Discussion ... 103
4.4.1. Variation in chlorophyll concentration ... 103
4.4.2. Predictive models ... 104
4.5. Conclusions ... 106
4.6. Acknowledgements ... 106
4.7. References ... 107
Chapter 5 Conclusions ... 112
5.1. Summary and findings ... 112
5.2. Limitations and future directions ... 116

List of Figures

Figure 1.1. (a) Color infrared Hyperion image of the Groundhog River Fluxnet site (GRFS) (shown by the white circle) and the vicinity area. (b) The three-dimensional hyperspectral data cube of the same image. The z-direction shows reflectance across 155 bands. Red and blue colors indicate high and low reflectance, respectively. (c) The spectral reflectance curve of a small area composed of conifer and deciduous species within the GRFS across 155 bands in the visible and near infrared regions of the electromagnetic spectrum. Reflectance is scaled by a factor of 10000. ... 6

Figure 1.2. LiDAR point cloud profile showing the vertical and horizontal structure of a mixedwood deciduous forest patch in northern Maryland. ... 10

Figure 2.1. (a) Location of the GRFS in Ontario, (b) Plot layout and biochemical sampling plot locations at the mixedwood forest site. WS = white spruce (*Picea glauca*); WC = northern white cedar (*Thuja occidentalis*); BS = black spruce (*Picea mariana*); WB = white birch (*Betula papyrifera*); TA = trembling aspen (*Populus tremuloides*); BF = balsam fir (*Abies balsamea*). . 26

Figure 2.2. Spatial distribution of the canopy N:P ratio at GRFS in 2004 and 2005 based on the predictive models canopy N:P 2004 = $9.19 + 0.73*(D_{max(680-750)}) - 12.59*(R_{734}-R_{747})/(R_{715}+R_{726})$, and canopy N:P 2005 = $13.78 + 0.79*(R_{750}/R_{700}) - 2.14*(R_{554}/R_{677})$. R and D represent reflectance and derivative of reflectance, respectively. ... 39

Figure 3.1. (a) Location of the GRFS, (b) Flux tower footprint at the mixedwood forest at the GRFS, (c) Plot layout and sample plot locations at the mixedwood forest site. WS = white spruce (*Picea glauca*); WC = northern white cedar (*Thuja occidentalis*); BS = black spruce (*Picea mariana*); WB = white birch (*Betula papyrifera*); TA = trembling aspen (*Populus tremuloides*); BF = balsam fir (*Abies balsamea*). ... 63

Figure 3.2. The relationship between: (a) Ca concentration and dominant height, and (b) N concentration and crown closure. W stands for the Shapiro-Wilk test statistic. BF=balsam fir, BS=black spruce, mix1=mixed deciduous, mix2=mixed deciduous and conifer, mix3=mixed conifer, TA=trembling aspen, WB=white birch, WC=white cedar. 71

Figure 3.3. (a) Observed versus predicted K concentration obtained from the IS and LiDAR model in Table 2c with the equation K = -606 + 247*$R_{1679\ nm}$ + 162*50th perc. ht. W stands for the Shapiro-Wilk test statistic. BF=balsam fir, BS=black spruce, mix1=mixed deciduous, mix2=mixed deciduous and conifer, mix3=mixed conifer, TA=trembling aspen, WB=white birch, WC=white cedar. (b) Spatial distribution of canopy K concentration at the mixedwood forest site calculated from this model. 74

Figure 4.1. Location of the Groundhog River Fluxnet Site (GRFS) and Sudbury in Ontario, Canada. Plots were established within the 1-km footprint of the flux tower at GRFS and just north of Sudbury. 90

Figure 4.2. Observed versus model predicted chlorophyll concentration using DCI and $D_{max(680-750)}$)/D_{703} indices with the same model parameters as generated from 2005 data (-8.2+28.1*(DCI) & -27.3+39* ($D_{max(680-750)}$)/D_{703}). The top graphs show 2005 model predictions for 2002, the middle ones for 2004 and the bottom ones for 2008 chlorophyll concentration, respectively. The prediction RMSEs are reported as percent of the mean of the chlorophyll concentration. BF=balsam fir, BS=black spruce, JP=jack pine, mix1=mixed deciduous, mix2=mixed deciduous and conifer, mix3=mixed conifer, TA=trembling aspen, WB=white birch, WC=white cedar. . 101

Figure 5.5. Scatterplot of foliar N, P and K samples of the six species found at the GRFS (color coded) in 3-D space. 116

List of Tables

Table 1.1. The estimated area and carbon stocks of major forest biomes (Dixon et. al., 1994)..... 3

Table 2.1. Foliar N:P ratio values for species found at GRFS, Ontario, Canada for 2004 and 2005. Values are mean ± standard deviation. n=5 for each species. * indicates N:P ratio differs significantly between years. 34

Table 2.2. N:P ratio predictive model statistics generated from 2004 airborne and 2005 spaceborne IS data. The robustness of the indices as predictors of the N:P ratio across time and scale is tested by applying the high resolution indices derived from the 2004 CASI data to the coarse resolution 2005 Hyperion data (High → Low) and vice versa (Low → High). R and D represent reflectance and derivative of reflectance, respectively. 35

Table 2.3. Effect of variation in scale and estimation method (i.e. map division vs. regression) on the canopy N:P ratio prediction accuracy in 2004 and 2005. 36

Table 2.4. LiDAR predictive model statistics for the canopy N:P ratio estimation. 37

Table 2.5. Correlation coefficients between canopy biophysical variables and the canopy N:P ratio. Statistical significance at 95% confidence or better is indicated by *..... 37

Table 2.6. Crown closure values of species at the plot level. Values are mean ± standard deviation. Lowercase letters denote significant differences. 38

Table 3.1. Site and species average macronutrient concentrations (mean ± standard deviation) for the mixedwood site at the Groundhog River Flux Station, 2005 (n=5 for all species except *Picea mariana* whose n=13). 68

Table 3.2. Predictive models generated from: (a) IS reflectance data, (b) LiDAR data, and (c) IS and LiDAR data combined. 69

Table 3.3. Spearman's correlation coefficients between macronutrient concentration and canopy structural metrics (* and ** denote significance at p<0.01 and p<0.0001, respectively)..... 70

Table 3.4. Spearman's correlation coefficients between canopy structural metrics and LiDAR metrics selected for the predictive models (* and ** denote significance of p<0.01 and p<0.0001, respectively)..... 72

Table 3.5. Spearman's correlation coefficients among the macronutrient concentrations of foliar samples collected at the Groundhog River Flux Site in July 2005 (* and ** denote significance at p<0.05 and p<0.0001, respectively)..... 73

Table 4.1. A subset of the spectral indices used for chlorophyll prediction. The name of the index, its formula and the reference when it was first developed is listed for each index. 94

Table 4.2. Average foliar chlorophyll concentrations ($\mu g/cm^2$) of individual species and all trees. Values are reported as mean ± standard deviation followed by the number of trees sampled for the analysis for a given species in parentheses. Sample size equals 5 for all other species. Uppercase letters denote significant differences in average foliar chlorophyll concentration for a given site (the first row) and for each species among years and lowercase letters denote significant differences in chlorophyll concentration among species within each year at $p \leq 0.05$, respectively. 98

Table 4.3. The performance of DCI and $D_{max(680-750)})/D_{703}$ for predicting chlorophyll concentration. PRESS RMSE and PRESS r^2 represent leave-one-out cross-validation error and coefficient of determination, respectively. * implies that the residuals fail to meet the normal distribution assumption..... 100

Table 4.4. The statistics associated with the five best two-variable models generated from 2005 data. PRESS RMSE and PRESS r^2 represent leave-one-out cross-validation error and coefficient of determination, respectively..... 102

Table 4.5. The predictions of chlorophyll concentration of 2002, 2004 and 2008 by the best two-variable models with the same model parameters as generated from 2005 data. For each model and year, r^2, adjusted r^2 and the RMSE (reported as percent of the mean of the chlorophyll concentration) are listed in the same order, respectively. 103

Chapter 1 General Introduction and Objectives

1.1. Introduction

1.1.1. Importance of Foliar Biochemicals

The biochemical components of a leaf can be divided into two major groups as pigment and non-pigment. Pigments are composed of chlorophylls, carotenoids, xanthopylls. Chlorophylls are the most important of pigments. Major non-pigment biochemicals are proteins, macronutrients, lignin, and cellulose. Nitrogen (N), phosphorus (P), potassium (K), magnesium (Mg) and calcium (Ca) constitute the primary macronutrients in plants. Measurement of both foliar chlorophyll and foliar macronutrients provide critical information about plant physiological and nutritional status, stress, as well as ecosystem processes such as carbon (C) exchange (photosynthesis and net primary production), decomposition and nutrient cycling. For example, N is an important constituent of the chlorophyll molecule and the carbon-fixing enzyme ribulose-1,5-bis-phosphate carboxylase/oxygenase and is thus directly related to photosynthesis (Field and Mooney, 1986). Foliar N is also related to primary production and decomposition (Melillo et al., 1982; Smith et al., 2002). In addition, the ratio of N to P (N:P ratio) in vegetation has been used as an index to detect nutrient limitation (Koerselman and Meuleman, 1996). P is a component of nucleic acids, lipid membranes, sugar phosphates and ATP which all have important roles in photosynthesis and respiration (Taiz and Zeiger, 2010). N and P are the major growth-limiting nutrients for plants worldwide (Chapin, 1980). Mg, like N, is also a constituent of the chlorophyll molecule and is directly related to photosynthesis (Taiz and Zeiger, 2010). K has a number of important roles in photosynthesis and respiration, including translocation of photosynthates into sink organs, maintenance of turgor pressure,

activation of enzymes, N metabolism and reducing excess uptake of ions such as Na and Fe in saline and flooded soils (Marschner, 1995; Mengel and Kirkby, 2001). Ca is required during cell division and in the synthesis of new cell walls, particularly the middle lamellae (Taiz and Zeiger, 2010).

Knowing foliar chlorophyll concentration is important because it is an indicator of stress, health and nutritional status of plants. There are several reasons for this. First, the amount of solar radiation absorbed by a leaf is largely depends on the foliar concentrations of chlorophyll, and therefore low concentration of chlorophyll can directly limit photosynthetic potential and hence primary production (Curran et al., 1990; Filella et al., 1995). Second, there is a close relationship between chlorophyll and leaf N since much of leaf N is incorporated in chlorophyll, both in its molecular structure and in the carbon-fixing enzyme ribulose-1,5-bis-phosphate carboxylase/oxygenase (Rubisco) so measurement of chlorophyll provides an indirect estimation of nutrient status (Filella et al., 1995; Moran et al., 2000). Third, concentration of chlorophyll is linked to stress where concentration of chlorophylls generally decreases under stress and during senescence (Peñuelas and Filella, 1998). Fourth, the relative concentrations of chlorophyll pigments a and b change based on abiotic factors such as light (e.g. sun leaves have a higher Chl a: Chl b ratio; Larcher, 1995). Therefore, quantifying the proportions can provide important information about relationship between plants and their abiotic environment.

1.1.2. The importance of the boreal forest for canopy biochemistry analysis

The boreal forest forms a circumpolar belt between the tundra in the north and the temperate forests and grasslands to the south (Larsen, 1980). It is mainly composed of coniferous fir, spruce, pine, larch, and deciduous aspen and birch species. Boreal forests influence the earth's climate through their effects on the atmospheric levels of CO_2, which is a greenhouse gas.

Even though they only approximately 33% of total forest area, they contain almost half of the C in the world's forests. C is stored in live plants in addition to the substantial accumulations in the soil over the long-term (Table 1.1) (Dixon et al., 1994).

Table 1.1. The estimated area and carbon stocks of major forest biomes (Dixon et. al., 1994).

Forest biome	Area (Mha)	Carbon pool Vegetation (Gt)	Carbon pool Soil (Gt)
boreal	1372	88	471
temperate	1038	59	100
tropical	1755	212	216

Warming trends would stimulate the decomposition of this substantial amount of stored soil C of the boreal forests and release into the atmosphere as CO_2, potentially transforming boreal forests from a C sink to a C source. Subsequently, this could exacerbate the impact of climate change caused by increased levels of CO_2 and other greenhouse gases. Because of their importance in the global C cycle, boreal forests have received a lot research attention.

Many studies have focused on trying to understand how these ecosystems would respond to increasing atmospheric concentrations of CO_2 and warming temperatures. For example, the Boreal Ecosystem–Atmosphere Study (BOREAS), a large-scale, international and interdisciplinary experiment with the goal of understanding how boreal forests interact with the atmosphere, how much CO_2 they are capable of storing, and how climate change will affect them was initiated in the northern boreal forests of Canada in 1993 (Sellers et al., 1997). Because of the boreal forests large extent and effects on global climate, ecological process models that provide estimations of productivity for large areas and allow for simulating different scenarios of climate conditions (e.g. doubling atmospheric concentration of CO_2 or increasing temperatures by several degrees) have been utilized (Kimball et al., 1997; 2000). These models may yield more accurate predictions when their input parameters are more detailed informational layers

such as remote sensing maps as opposed to broad average values. In this regard, landscape level maps of canopy macronutrients, in particular N, because of its close relationship to primary productivity (Smith et al., 2002) may provide very useful to improve the accuracy of the predictions of these models (Ollinger and Smith, 2005). Canopy chlorophyll has also been used to predict GPP of boreal forests coupled with satellite remote sensing data (Zhang et al., 2009). This would help scientists to argue their case towards the public and policy makers based on more sound results.

1.1.3. Remote Sensing of Foliar Biochemicals

1.1.3.1. Strategic Use of Remote Sensing in Ecology

The effects of climate change, land use change, and disturbance on the ecosystems of the world have resulted in an increased demand for remote sensing data at all scales. A wide range of information is needed to predict the consequences of climate change and to monitor C, water, and nutrient cycles, from land cover, land-use history, and estimates of standing biomass to succession, biodiversity, and sustainability. Traditional field-based sampling methods are prohibitively expensive and time-consuming at large spatial scales. Remote observations either from space or air provide practical means to obtain a synoptic view of Earth's ecosystems such as spatial distribution, extent and temporal dynamics at varying details (Cohen and Goward, 2004). Imaging spectroscopy (IS) and Light Detection and Ranging (LiDAR) are two remote technologies which provide information about physiology and structure of ecosystems, respectively.

1.1.3.2. Importance of Remotely Sensing Foliar Biochemicals at the Canopy Level

The remote detection of macronutrients in the canopy of forest ecosystems is important because it can be used to estimate important physiological and ecosystem processes such as photosynthesis and net primary production. For example, foliar N has been used as a parameter in ecosystem models (Parton et al., 1995; Ollinger and Smith, 2005) and canopy N concentration has been used to estimate net primary productivity of temperate forest ecosystems coupled with field measurements and hyperspectral remote sensing data (Smith et al., 2002). Thus, spatially-explicit landscape and regional inputs of foliar N concentration have the potential to improve the accuracy of ecosystem models. Maps of canopy macronutrient concentrations at small scales (e.g., in close proximity to a flux tower) can also serve as validation products for large-scale modeling efforts at continental and global scales. In addition, from a forest management perspective, spatial maps of canopy macronutrients could be used as indicators of nutrient deficiencies and consequently utilized to adjust the fertilizer scheme to optimize maximum yield with minimum cost.

1.1.3.3. Imaging Spectroscopy and Light Detection and Ranging for Canopy Biochemistry Estimation

Imaging spectroscopy proved to be very useful for canopy biochemistry estimation. One of the great advantages is that it allows for the remote detection of foliar biochemicals due to its high spectral resolution with many and narrow bands resulting in an almost continuous reflectance spectrum throughout the visible, near and shortwave-infrared portions of the electromagnetic spectrum (Goetz et al., 1985) (Figure 1.1).The specific absorption features attributable to certain chemicals at particular wavelengths can be detected and used as predictor variables to estimate the chemical of interest using regression analysis (Curran, 1989).

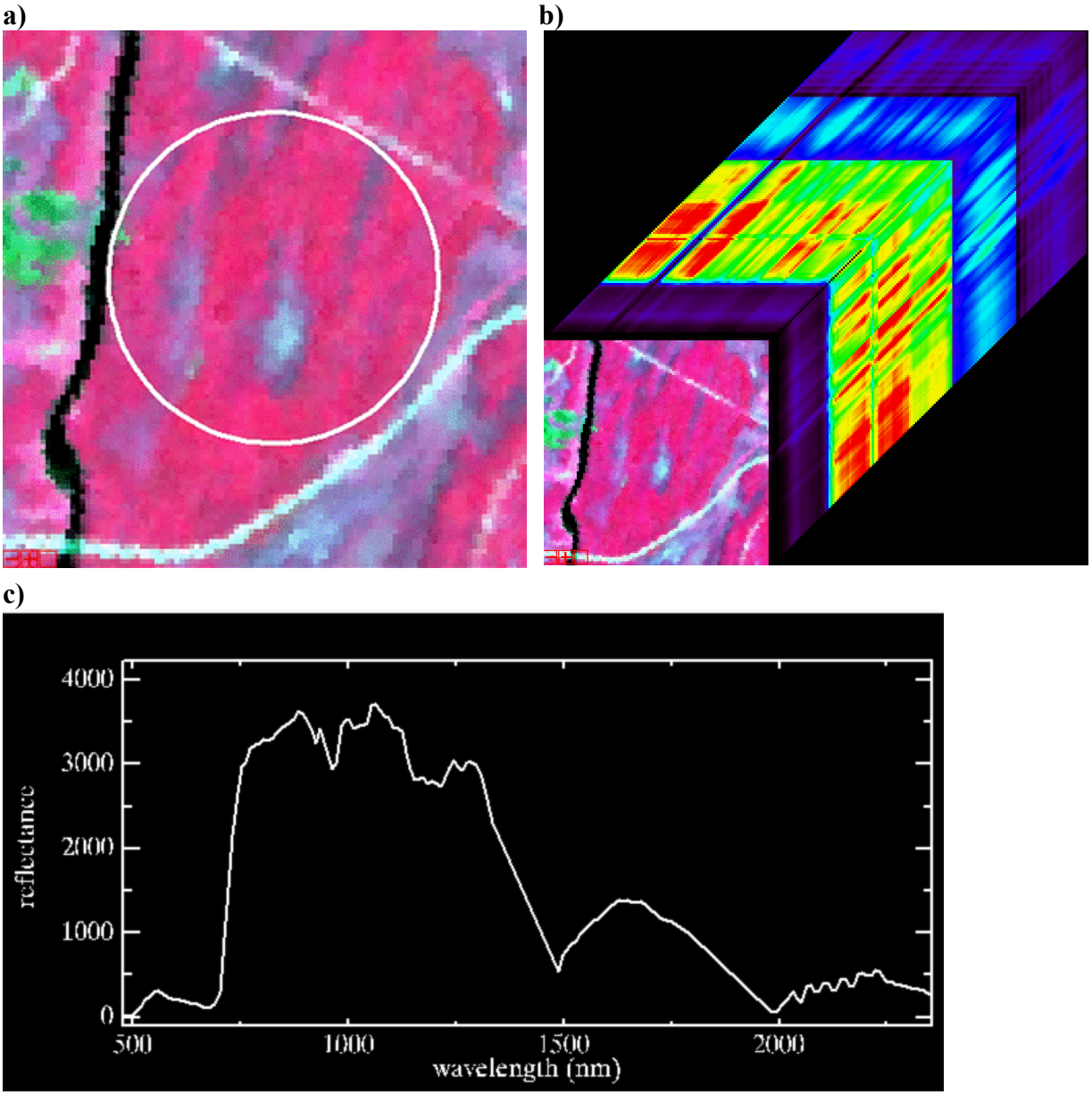


Figure 1.1. (a) Color infrared Hyperion image of the Groundhog River Fluxnet site (GRFS) (shown by the white circle) and the vicinity area. (b) The three-dimensional hyperspectral data cube of the same image. The z-direction shows reflectance across 155 bands. Red and blue colors indicate high and low reflectance, respectively. (c) The spectral reflectance curve of a small area composed of conifer and deciduous species within the GRFS across 155 bands in the visible and near infrared regions of the electromagnetic spectrum. Reflectance is scaled by a factor of 10000.

The remote sensing of foliar biochemicals using spectroscopy can be traced back to the 1970s when biochemicals in dried and ground leaves of agricultural forage crops were measured using near infrared spectroscopy in laboratory settings (Norris et al., 1976; Shenk et al., 1979) with support from U.S. Department of Agriculture (USDA). As a result of numerous studies, several absorption features related to the concentration of biochemicals including cellulose, lignin, protein, oil, sugar, starch and water were identified and reflectance spectroscopy became a routine procedure for the estimation of protein, lignin and starch concentrations in dried plant foliage in the USDA labs (Norris and Barnes, 1976; Marten et al., 1985; 1989; Weyer 1985; Williams et al., 1984; Williams and Norris, 1987). By the late 1980s, the procedures developed by the USDA to estimate chemicals using spectroscopy have been applied to estimate biochemical composition of dried and ground forest tree leaves (Card et al., 1988; Wessman et al., 1988a) and fresh conifer leaves (Peterson et al., 1988) in the laboratory. Finally, with the advent of airborne imaging spectrometer (AIS) (Goetz et al., 1985), forest canopy biochemistry in the field was estimated (Peterson et al., 1988; Wessman et al., 1988b). Since then, numerous imaging spectrometers such as the airborne National Aeronautics and Space Administration (NASA) Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), Australian HyMap, Canadian Compact Airborne Imaging Spectrometer (CASI) and spaceborne EO-1 Hyperion have been developed. The data from these sensors have been successfully used to estimate foliar N, lignin and pigments in a variety of forest ecosystems (Wessman et al., 1988b; Zagolski et al., 1996; Curran et al., 1997; Martin and Aber, 1997; Coops et al., 2003; Thomas et al., 2008a). Some studies also compared the effectiveness of different sensor types, such as AVIRIS and Hyperion to estimate N concentration in temperate forest canopies (Smith et al., 2003; Townsend et al., 2003). More recently, P has also been estimated using imaging spectrometry. Studies

include using field spectrometers to predict P (Al-Abbas et al., 1974; Osborne et al., 2002; Gong et al., 2002) and examining the correlation of P with canopy reflectance (Asner et al., 2008). Landscape level estimation of P was carried out in African savanna using a HyMap image (Mutanga and Kumar, 2007) and in Hawaiian tropical forest canopy using AVIRIS data (Porder et al., 2005). Spectral characteristics of other important plant nutrients including Mg, K, Ca, sodium and sulfur were studied using spectrometers in the lab and field (Al-Abbas et al., 1974; Mutanga et al., 2004; Ferwerda and Skidmore, 2007; Ponzoni and Gonçalves, 1999).

Hyperspectral narrowband, red-edge (approximately 680-750 nm range) derivative indices have been used to predict chlorophyll and other plant pigments (Curran et al., 1990; Gitelson and Merzlyak, 1996; Blackburn, 1998; Zarco-Tejada et al., 2000; Haboudane et al., 2002). Mutanga and Skidmore (2004) used narrowband vegetation indices to estimate biomass. Thenkabail et al. (2004) utilized hyperspectral indices for land cover classification. N concentration of wheat was predicted using NDVI-based indices by Hansen and Schoerring (2003). Ferwerda et al. (2005) examined the 300-2500 nm spectral range for N prediction using normalized ratio indices.

The robustness and predictive capabilities of broadband and hyperspectral vegetation indices for estimating canopy biophysical and biochemical properties using coupled leaf and canopy radiative transfer models was examined (Broge and Leblanc, 2000). They found that canopy architecture was an important factor to consider for testing the performance of the indices in addition to illumination and atmospheric conditions. The 3-D structural information of the vegetation was found to be important when leaf chlorophyll concentrations were estimated by Demarez and Gastellu-Etchegorry (2000). They also noted the computation to obtain the 3-D

structure information could be intensive depending on the canopy complexity. A remote sensing technique that is capable of providing canopy structural information is LiDAR.

LiDAR is an active remote sensing technique where laser pulses are usually sent off an airplane and the height of objects on Earth's surface are calculated using the return time of the pulse to the origin (van Leeuwen and Nieuwenhuis, 2010). The distribution of returns in three dimensional space can characterize the vertical and horizontal structure of the canopy and subcanopy (Figure 1.2). LiDAR data have been used to characterize the structure and biophysical properties of forest ecosystems (Magnussen and Boudewyn, 1998; Lefsky et al., 1999; Lim and Treitz, 2004; Popescu et al., 2003; Riaño et al., 2004). The general approach in these studies has been to calculate metrics related to height from the LiDAR returns such as mean height, maximum height and quantiles and then relating these LiDAR metrics to biophysical properties measured in the field such as mean tree height, maximum tree height, crown closure, volume, biomass and leaf area index (LAI) using statistical analyses, typically regression. A combined method where the spectral information from imaging spectroscopy and the structural information from the LiDAR data have been utilized to estimate chlorophyll in forest canopies (Blackburn, 2002; Thomas et al., 2008). These studies showed that the predictive accuracy of the models improved with the addition of LiDAR structural data.

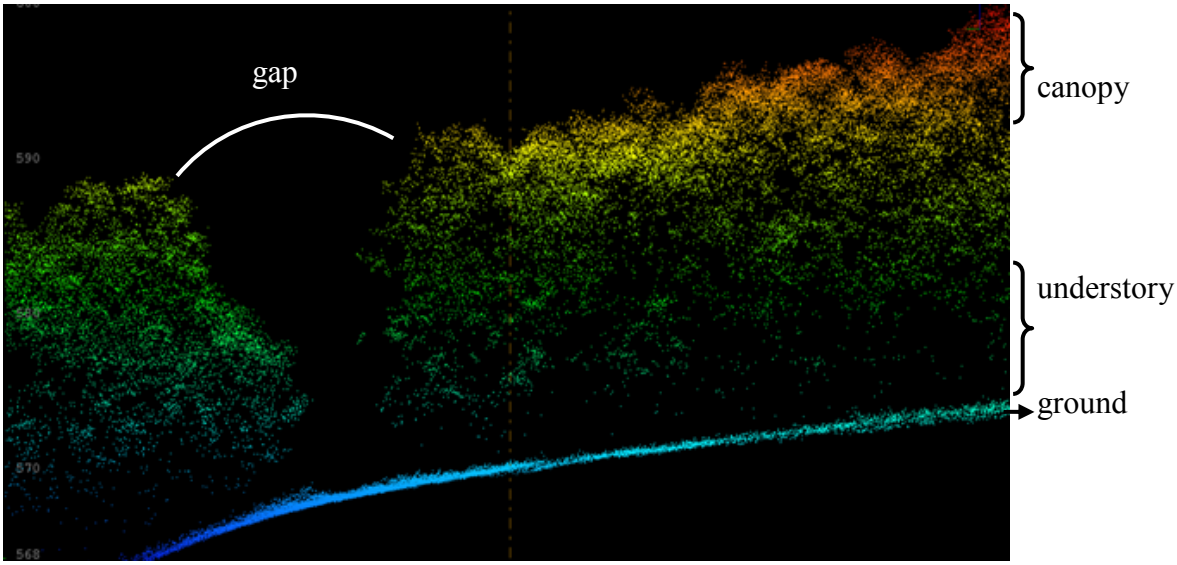


Figure 1.2. LiDAR point cloud profile showing the vertical and horizontal structure of a mixedwood deciduous forest patch in northern Maryland.

1.2. Objectives

The goal of this study was to explore the estimation of the N:P ratio, macronutrients (N, P, K, Ca and Mg) and chlorophyll pigment in the canopy of a mixedwood boreal forest using satellite IS data. Airborne LiDAR data were used as a supplementary dataset to test whether they would provide improvements for the predictions. The specific objectives are listed below and each constitutes an independent research chapter.

Objective 1: to estimate the canopy N:P ratio using airborne and spaceborne IS data, address the effects of temporal and spatial resolution variation on prediction accuracy of the canopy N:P ratio and investigate whether LiDAR data have any explanatory power for the estimation of the canopy N:P ratio. Chapter 2 addresses the possibility of remotely detecting nutrient limitation for biomass production in a boreal mixedwood community, which can be determined by the N:P

ratio. Both airborne and spaceborne IS data are used and the robustness of spectral indices from each dataset is tested.

Objective 2: to evaluate spaceborne IS data to estimate canopy macronutrient (N, P, K, Ca and Mg) concentrations in a mixedwood boreal forest; and test the potential contribution of canopy structural information derived from LiDAR data for improving the predictive models for these macronutrients. Objective 2 is addressed in Chapter 3.

Objective 3: to identify spectral indices that can predict canopy chlorophyll concentration across years, test the robustness of the models with these indices derived from spaceborne IS data for predicting canopy chlorophyll concentration across years within the same site, and extend these predictive models for canopy chlorophyll concentration to a different geographic location and time. These objectives were addressed in Chapter 4.

1.3. Importance of research

The remote detection of the N:P ratio provides a means to assess nutrient limitation of biomass production at local scales and brings the possibility of extending it to regional scales using spaceborne data if robust methods can be found (Chapter 2). If predictive models that are comparable in accuracy to the ones generated from airborne data can be obtained for the macronutrients, then it would be possible to extend the analysis at the canopy scale over larger geographic areas at a lower cost. The work is done in anticipation of future availability of global IS data from spaceborne platforms, such as the German Aerospace Center's European Environmental Mapping and Analysis Program (EnMAP), with a 2015 expected launch date, and the Hyperspectral Infrared Imager (HyspIRI) mission of NASA, which is expected to be

launched between 2013-2016 (DLR, 2012; NASA, 2012). The spatially-explicit maps of nutrient concentrations may be used as inputs into ecosystem productivity models (Chapter 3). In both of these chapters, LiDAR data are used to improve predictive models by linking them to forest structural attributes. This is different from the traditional way of extracting metrics from LiDAR data and establishing predictive models using these metrics as predictors to estimate forest structural characteristics such as height, biomass and LAI (van Leeuwen and Nieuwenhuis, 2010). The fourth chapter not only deals with using spaceborne IS data for predicting chlorophyll in a boreal forest canopy but also addresses the problem of the non-transferability of empirical models across time and space. It tests the robustness of these models for predicting canopy chlorophyll at different times and a different location with different sensor data. The findings from this chapter pave the way for a universal method that could be applied to the mixedwood boreal forest for predicting chlorophyll at the canopy scale.

1.4. References

Al-Abbas, A.H., Barr, R., Hall, J.D., Crane, F.L., Baumgardner, M.F., 1974. Spectra of normal and nutrient deficient maize leaves. Agronomy Journal 66, 16–20.

Asner, G.P., Jones, M.O., Martin R.E., Knapp, D.E., Hughes, R.F., 2008. Remote sensing of native and invasive species in Hawaiian forests. Rem. Sens. Environ. 112, 1912–1926.

Blackburn, G.A. 1998. Quantifying chlorophylls and carotenoids at leaf and canopy scales: an evaluation of some hyperspectral approaches. Rem. Sens. Environ. 66, 273–285.

Blackburn, G.A. 2002. Remote sensing of forest pigments using airborne imaging spectrometer and LIDAR imagery. Rem. Sens. Environ. 82, 311–321.

Broge, N.H., Leblanc, E., 2000. Comparing prediction power and stability of broadband and hyperspectral indices for estimation of green leaf area index and canopy chlorophyll density. Rem. Sens. Environ. 76, 156–172.

Card, D.H., Peterson, D.L., Matson, P.A., 1988. Predicting of leaf chemistry by the use of visible and near infrared reflectance spectroscopy. Rem. Sens. Environ. 26, 123-147.

Chapin, F S. 1980. The mineral nutrition of wild plants. *Annual Review of Ecology and Systematics*, Vol. 11, pp. 233-260.

Cohen, W.B., Goward, S.N. 2004. Landsat's Role in Ecological Applications of Remote Sensing. Bioscience, 54-6: 535-545.

Coops, N.C., Smith, M.L., Martin, M.E., Ollinger, S.V., 2003. Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data. IEEE Trans. Geosci. Rem. Sens. 41:6, 1338-1346.

Curran, P. J., 1989. Remote sensing of foliar chemistry. Rem. Sens. Environ. 30, 271−278.

Curran, J.P., Dungan, J.L., Gholz, H.L., 1990. Exploring the relationship between reflectance red edge and chlorophyll content in slash pine. Tree Physiol. 7, 33-48.

Curran, P.J., Kupiec, J.A., Smith, G.M., 1997. Remote sensing the biochemical composition of a slash pine canopy. IEEE Trans. Geosci. Rem. Sens. 35:2, 415-420.

Demarez, V. and Gastellu-Etchegorry, J.P. 2000. A Modeling Approach for Studying Forest Chlorophyll Content. Rem. Sens. Env., 71: 226-238.

Dixon, R.K., Brown, S., Houghton, R.A., Solomon, M.A., Trexler, M.C., Wisniewski, J. 1994. Carbon Pools and Flux of Global Forest Ecosystems. Science 263: 185-190.

DLR, 2012. German Aerospace Center. Available at the Earth Observation Center of DLR http://www.enmap.org/mission_statement (accessed on June 24, 2012).

Ferwerda, J.G., Skidmore, A.K., Mutanga, O., 2005. Nitrogen detection with hyperspectral normalized ratio indices across multiple plant species. Int. J. Rem. Sens. 26:18, 4083–4095.

Ferwerda, J.G. and A.K. Skidmore. 2007. Can nutrient status of four woody plant species be predicted using field spectrometry? ISPRS Journal of Photogrammetry and Remote Sensing, 62: 406-414.

Field, C., and Mooney, H.A.1986. The photosynthesis–nitrogen relationship in wild plants. In On the Economy of Plant Form and Function: Proceedings of the Sixth Maria Moors Cabot Symposium, Evolutionary Constraints on Primary Productivity, Adaptive Patterns of Energy Capture in Plants, August 1983, Harvard Forest, USA. Edited by T.J. Givnish. Cambridge University Press, Cambridge, pp. 25–55.

Filella, I., Serrano, L., Serra, J., Peñuelas, J. 1995. Evaluating wheat nitrogen status with canopy reflectance indices and discriminant analysis. Crop Science 35: 1400-1405.

Gitelson, A., Merzlyak, M.N., 1996. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. J. Plant Physiol. 148, 501–508.

Goetz, A.F.H., Vane, G., Solomon, J.E., and B.N. Rock. 1985. Imaging Spectrometry for Earth Remote Sensing. Science, 228-4704: 1147-1153.

Gong, P., Pu, R., Heald, R.C., 2002. Analysis of in situ hyperspectral data for nutrient estimation of giant sequoia. Int. J. Rem. Sens. 23:9, 1827–1850.

Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J. Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Rem. Sens. Environ. 81, 416–426.

Hansen, P.M., Schoerring, J.K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Rem. Sens. Environ. 86, 542–553.

Kimball, J.S., Thornton, P.E., White, M.A., Running, S.W. 1997. Simulating forest productivity and surface–atmosphere carbon exchange in the BOREAS study region. Tree Physiology 17: 589-599.

Kimball, J.S., Keyser, R.A., Running, S.W, Saatchi, S.S. 2000. Regional assessment of boreal forest productivity using an ecological process model and remote sensing parameter maps. Tree Physiology 20: 761-775.

Koerselman, W. and Meuleman, A.F.M. 1996. The vegetation N:P ratio: a new tool to detect the nature of nutrient limitation. *Journal of Applied Ecology* 33: 1441–1450.

Larcher, W. 1995. Physiological Plant Ecology. 3rd Ed. Berlin, Germany, Springer.

Larsen, J.A. 1980. The boreal ecosystem. Academic Press, New York.

Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., Harding, D., 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Rem. Sens. Environ. 70, 339–361.

Lim, K.S., and Treitz, P.M. 2004. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research*, Vol. 19, No. 6, pp. 558–570.

Magnussen, S., and Boudewyn, P. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research, 28: 1016-1031.

Marschner, H. 1995. *Mineral Nutrition of Higher Plants*. 2nd Edition, Academic Press, San Diego.

Marten, G. C., J. S. Shenk, and F. E. Barton, II. 1985. Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality (eds). Agric. Handbook No. 643.United States Department of Agriculture- Agricultural Research Service, Washington, DC.

Marten, G.C., Shenk, J.S., and Barton, F.E. II. 1989. Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality (eds). Agric. Handbook 643. United States Department of Agriculture- Agricultural Research Service, Washington, DC.

Martin, M. E., Aber, J.D., 1997. High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. Ecol. Appl. 7, 431−444.

Melillo, J. M., Aber, J.D., and Muratore, J.M. 1982. Nitrogen and lignin control of hardwood leaf litter decomposition dynamics. *Ecology*, Vol. 63, pp. 621–626.

Mengel, K., and Kirkby, E. A. 2001. *Principles of Plant Nutrition*. 5th Edition, Kluwer Academic Publishers, Dordrecht.

Moran, J.A., Mitchell, A.K, Goodmanson, G., Stockburger, K.A. 2000. Differentiation among effects of nitrogen fertilization treatments on conifer seedlings by foliar reflectance: a comparison of methods. Tree Physiology 20: 1113-1120.

Mutanga, O., Skidmore, A.K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. Int. J. Rem. Sens. 25:19, 3999–4014.

Mutanga, O., Skidmore, A.K., and H.H.T. Prins. 2004. Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. Remote Sensing of Environment, 89: 393–408.

Mutanga, O., Kumar, L., 2007. Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. Int. J. Rem. Sens. 28:21, 4897–4911.

NASA, 2012. HyspIRI mission website available at http://hyspiri.jpl.nasa.gov (accessed on June 24, 2012.

Norris, K. H., and Barnes, R. F. 1976. Infrared reflectance analysis of nutrient value of feed-stuff, in Proceedings, 1st International Symposium, Feed Composition. Animal Nutrient Requirements and Computerization of Diets', Utah State University, Logan, UT, pp. 237-241.

Norris, K.H., Barnes, R.F., Moore, J.E., Shenk, J.S., 1976. Predicting forage quality by infrared reflectance spectroscopy. J. Anim. Sci. 43, 889-897.

Ollinger, S.V., and M.-L. Smith, 2005. Net primary production and canopy nitrogen in a temperate forest landscape: An analysis using imaging spectroscopy, modeling, and field data, Ecosystems, 8:1–19.

Osborne, S.L., Schepers, J.S., Francis, D.D., Schlemmer, M.R., 2002. Detection of phosphorous and nitrogen deficiencies in corn using spectral radiance measurements. Agron. J. 94, 1215–1221.

Parton, W.J., Scurlock, J.M.O., Ojima, D.S., Schimel, D.S., Hall, D.O., Scopegram Group Members. 1995. Impact of climate change on grassland production and soil carbon worldwide. Global Change Biology, 1: 13-22.

Peñuelas, J., Filella, I. 1998. Visible and near-infrared reflectance techniques for diagnosing plant physiological status. Trends in Plant Science 3: 151-156.

Peterson, D.L., Aber, J.D., Matson, P.A., Card, D.H., Swanberg, N., Wessman, C., Spanner, M., 1988. Remote sensing of forest canopy and leaf biochemical contents. Rem. Sens. Environ. 24, 85-108.

Ponzoni, F.J., Gonçalves, J.L. de M., 1999. Spectral features associated with nitrogen, phosphorus, and potassium deficiencies in Eucalyptus saligna seedling leaves. Int. J. Rem. Sens. 20:11, 2249-2264.

Popescu SC, Wynne RH, Nelson RF. 2003. Measuring individual tree crown diameter with LiDAR and assessing its influence on estimating forest volume and biomass. Can J Remote Sens 29: 564–577.

Porder, S., Asner, G.P., Vitousek, P.M., 2005. Ground-based and remotely sensed nutrient availability across a tropical landscape. Proc. Natl. Acad. Sci. U.S.A. 102:31, 10909–10912.

Riaño, D., Valladares, F., Condés, S., and Chuvieco, E. 2004. Estimation of leaf area index and covered ground from airborne laser scanner (Lidar) in two contrasting forests. Agricultural and Forest Meteorology, 124: 269-275.

Sellers, P.J., F.G. Hall, R.D. Kelly, A. Black, D. Baldocchi, J. Berry, M. Ryan, K.J. Ranson, P.M. Crill, D.P. Lettenmaier, H. Margolis, J. Cihlar, J. Newcomer, D. Fitzjarrald, P.G. Jarvis, S.T. Gower, D. Halliwell, D. Williams, B. Goodison, D.E. Wickland and F.E. Guertin. 1997. BOREAS in 1997: experiment overview, scientific results, and future directions. J. Geophys. Res. 102: 28731–28769.

Shenk, J.S., Westerhous, M.O., Hoover, M.R., 1979. Analysis of forages by infrared reflectance. J. Dairy Sci. 62, 807−812.

Smith, M.L., Ollinger, S.V., Martin, M.E., Aber, J.D., Hallett, R.A., Goodale, C.L., 2002. Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. Ecol. Appl. 12, 1286−1302.

Smith, M.L., Martin, M.E., Ollinger, S.V., Plourde, L., 2003. Analysis of hyperspectral data for estimation of temperate forest canopy nitrogen concentration: Comparison between an airborne (AVIRIS) and a spaceborne (Hyperion) sensor. IEEE Trans. Geosci. Rem. Sens. 41:6, 1332−1337.

Taiz, L., and Zeiger, E. 2010. *Plant Physiology*. 5th Edition, Sinauer Associates Inc., Sunderland. Thomas, V., Treitz, P., McCaughey, J.H., Noland, T., Rich, L., 2008a. Canopy chlorophyll concentration estimation using hyperspectral and lidar data for a boreal mixedwood forest in northern Ontario, Canada. Int. J. Remote Sens. 29, 1029–1052.

Thenkabail, P.S., Enclona, E.A., Ashton, M.S., van der Meer, B., 2004. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. Rem. Sens. Environ. 91, 354–376.

Thomas, V., Treitz, P., McCaughey, J.H., Noland, T., Rich, L., 2008. Canopy chlorophyll concentration estimation using hyperspectral and lidar data for a boreal mixedwood forest in northern Ontario, Canada. Int. J. Remote Sens. 29, 1029–1052.

Townsend, P.A., Foster, J.R., Chastain, Jr., R.A., Currie, W.S., 2003. Imaging spectroscopy and canopy nitrogen: Application to the forests of the central Appalachian Mountains using Hyperion and AVIRIS. IEEE Trans. Geosci. Remote Sens. 41:6, 1347−1354.

van Leeuwen, M., and Nieuwenhuis, M. 2010. Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research*, Vol. 129, No.4, pp. 749–770.

Wessman, C.A., Aber, J.D., Peterson, D.L., Melillo, J.M., 1988a. Foliar analysis using near infrared reflectance spectroscopy. Can. J. For. Res. 18, 6-11.

Wessman, C.A., Aber, J.D., Peterson, D.L., Melillo, J.M., 1988b. Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. Nature 335, 154-156.

Weyer, L.G. 1985. Near infrared spectroscopy of organic substances. Appl. Spectrosc. Rev. 21:1-43.

Williams, P., and Norris, K., Eds. 1987. Near-Infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemists, St. Paul, MN.

Williams, P.C., Preston, K.R., Norris, K.H., Starkey, P.M., 1984. Determination of amino acids in wheat and barley by near-infrared reflectance spectroscopy. J. Food Sci. 49, 17−20.

Zagolski, F., V. Pinel, J. Romier, D. Alcayde, J. P. Gastellu- Etchegorry, G. Giordano, G. Marty, and E. Mougin. 1996. Forest canopy chemistry with high spectral resolution remote sensing. International Journal of Remote Sensing 17: 1107–1128.

Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammad, G.H., Sampson, P.H., 2000. Chlorophyll fluorescence effects on vegetation apparent reflectance:II Laboratory and airborne canopy-level measurements with hyperspectral data. Rem. Sens. Environ. 74, 596–608.

Chapter 2 Canopy nitrogen to phosphorus ratio mapping at two different scales in a boreal mixedwood forest

K. Gökkaya [a], V. Thomas [a], T. Noland [b], J.H. McCaughey [c], I. Morrison [d] and P.M. Treitz [c]

[a] Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, USA
[b] Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, ON, Canada
[c] Department of Geography, Queen's University, Kingston, ON, Canada
[d] Canadian Forest Service, Natural Resources Canada, Sault Ste. Marie, ON, Canada

This chapter is under review at Applied Vegetation Science.

Abstract

The ratio of nitrogen to phosphorus (N:P ratio) in vegetation has been used as an index to detect nutrient limitation of biomass production and nutrient availability in plant communities. In this study, we explored the possibility of the estimation of the spatial pattern of foliar N:P ratio using imaging spectroscopy (IS) data at two scales, airborne and spaceborne, in a boreal mixedwood canopy composed of black and white spruce, balsam fir, northern white cedar, white birch, and trembling aspen across two summers. The relationship between the canopy N:P ratio and forest structure was also investigated through the analysis of light detection and ranging (LiDAR) data. Airborne and spaceborne IS data explained 70 and 69% of the variance in canopy N:P ratio with predictions errors of 5 and 7.2% in two consecutive years, respectively. Airborne IS data were temporally and spatially more robust than spaceborne IS data in predicting the canopy N:P ratio and predictions differed significantly with changes in scale. Difference in spatial resolution has an impact on model performance for the canopy N:P ratio prediction as related to the modifiable areal unit problem (MAUP). Although the predictive models obtained from LiDAR data did not provide improved explanatory power or accuracy for canopy N:P ratio relative to the IS models (R^2 of 0.54 and 0.67, prediction errors of 6.1 and 7.5% for the two years, respectively), they provide insight into the relationship between structure and growth and productivity at the site. We suggest that the relationship between the canopy N:P ratio and remote sensing data at this site is based on the relationship between the canopy N:P ratio and crown closure. Presence of multiple species creates a gradient in the N:P ratio, making it feasible to predict the N:P ratio using remote sensing data. Results indicate that the canopy N:P ratio can be mapped using both airborne and spaceborne IS data at this site but variations in spatial and

temporal scales should be taken into account. Our results suggest that a canopy N:P ratio map has potential diagnostic value in determining nutrient limitations and its accuracy should be tested with the availability of data from future studies that focus on experimental work to identify the critical limiting values of the N:P ratio for biomass production in boreal ecosystems.

2.1. Introduction

Nitrogen (N) and phosphorus (P) are the major growth-limiting nutrients for plants worldwide (Chapin 1980). The ratio of nitrogen to phosphorus (N:P ratio) in vegetation has been used as an index to detect nutrient limitation of biomass production and nutrient availability in plant communities. This approach is well established in the discipline of ecological stoichiometry, which studies the balance of chemical elements in ecological processes and interactions (Elser et al. 2000). A wide range of N:P ratio values have been reported as thresholds for nutrient limitation of biomass production in wetlands (Koerselman & Meuleman 1996), grasslands (Craine et al. 2008), understory vegetation of temperate forests (Tessier & Raynal 2003), montane forests (Herbert & Fownes 1995), loblolly pine (*Pinus taeda* L.) plantations (Valentine & Allen 1990) and alpine meadow (Bowman 1994). In boreal forests, studies have been conducted in Norway spruce (*Picea abies* (L.) Karst.) and Scots pine (*Pinus sylvestris* L.) stands (Jacobson & Pettersson 2001), and Norway spruce plantations (Clarholm & Rosengren-Brinck 1995) in Europe, in black spruce (*Picea mariana* [Mill.] B.S.P.) forests of Quebec, Canada (Paquin et al. 1998) and jack pine (*Pinus banksiana* Lamb.) stands of northern Ontario, Canada (Foster & Morrison 1976). In addition, vegetation N:P ratios have been used to indicate N saturation in forests of western and eastern USA (Fenn et al. 1996; Tessier & Raynal 2003). Similarly, Gradowski & Thomas (2008) reported a high foliar N:P ratio for a sugar maple forest in Canada that was receiving high N deposition and where growth appeared to be P

limited. Recently, the variation in foliar N:P ratio in relation to biotic (i.e. functional groups) and abiotic factors (i.e. climate, topography, geographic location) across national, continental and global scales have been investigated (Reich & Oleksyn 2004; Han et al. 2005; Kang et al. 2011).

The importance of and interest in the foliar N:P ratio and its spatial variation suggests a need to analyze its spatial distribution. If correlations exist between remote sensing data and the N:P ratio, canopy level estimates of the N:P ratio can be achieved using remote sensing. A foliar N:P ratio map displaying its spatial variation at the canopy level can facilitate the diagnosis of nutrient availability and limitation and help to eliminate the need to do intensive field sampling or nutrient addition experiments. In addition, a map of the foliar N:P ratio generated from remote sensing data at a small scale may be used as a validation tool for larger scale modeling attempts at continental and global scales, which may become possible with the launch of satellites within this decade such as the HyspIRI and EnMAP, which will acquire global high spectral resolution data. Relative to multispectral sensors such as Landsat, imaging spectroscopy (IS), provides high spectral resolution as a result of contiguous and narrow spectral bands, which cover the visible and near infrared (VNIR) and shortwave infrared (SWIR) regions of the electromagnetic (EM) spectrum. Improved spectral resolution makes it possible to estimate certain biochemicals with specific absorption features within the EM spectrum by establishing relationships between leaf measurements and spectral information (Curran 1989). Airborne IS data have been used to estimate N concentration in the canopies of temperate forests (Wessman et al. 1988; Martin & Aber 1997) and African savannah grass (Mutanga & Skidmore 2004), spaceborne Hyperion IS data have been used to predict the N concentration of eucalyptus canopy (Coops et al. 2003). Airborne AVIRIS and HyMap IS data were used for estimating P concentrations in Hawaiian tropical forest and African savannah grass canopies, respectively (Porder et al. 2005; Mutanga &

Kumar 2007). Nitrogen and P were predicted together in the same location in the forage vegetation of Yellowstone National Park using airborne IS data (Mirik et al. 2005) and African savannah grass with spectroradiometer data (Mutanga et al. 2004). The relationship of N and P concentration to leaf and modeled canopy reflectance was investigated in Australian tropical forests by Asner & Martin (2008). Hyperion data were used to generate predictive models for N and P concentration in the boreal mixedwood forest of northern Ontario, Canada (Gökkaya et al. 2012, in review).

Spaceborne and airborne data usually have different spatial resolutions, with airborne remote sensing most often having higher spatial resolutions. The different spatial resolutions of remote sensing data are an inherent characteristic resulting from differences in sensors and platforms. Remote sensing data can be viewed as an arbitrary sampling grid superimposed over the surface of the earth. In fact, Marceau (1992) identified remote sensing data as a unique case of the modifiable areal unit problem (MAUP). In landscape ecology, a study area can be divided into non-overlapping areal units which can take any size and shape which affects the statistical relationships and models in two significant ways. First is the 'scale effect' attributed to variation in numerical results when areal units used in the analysis are progressively aggregated into fewer, larger units. Second issue is the zonation, or the 'aggregation effect' due to changes in numerical results as a result of alternative combinations of areal units at equal or similar resolutions (Openshaw & Taylor 1979). The scale effect is directly related to the remote sensing data because remote sensing imagery represents two aspects of scale: grain and extent. Grain corresponds to spatial resolution and extent represents the total area covered within an image (O'Neill et al. 1996). In remote sensing, the modifiable units are the pixels, and when equal-sized pixels are aggregated into fewer and larger units, the variation in the data decreases which

affects the results of statistical analysis (Jelinski & Wu 1996). As a result, any attempt to transfer models from remote sensing data at a given spatial resolution to a different resolution is likely to produce inaccurate results. Model parameters would need to be recalibrated beforehand. The influence of varying spatial resolution on classification accuracy (Arbia et al. 1996), detecting landscape patterns (Benson & Mackenzie 1995; Moody & Woodcock 1995; Pax-Lenney & Woodcock 1997) and forest process modeling (Turner et al. 1996; McNulty et al. 1997) have been investigated.

Since the N:P ratio is an index of nutrient limitation of biomass production and nutrient availability, it is related to primary productivity, manifested as biomass and in turn, tree and canopy height. Boreal mixedwood forests are composed of coniferous and deciduous species that have different heights and canopy shapes, contributing to structural complexity in three dimensions. This structural complexity can be characterized by light detection and ranging (LiDAR) remote sensing technology, which has the ability to penetrate the canopy and provide information about its vertical and horizontal structure. This allows one to detect any correlations between canopy height, structure and the foliar N:P ratio if indeed, they exist. LiDAR and IS data have been used to investigate boreal forest structure and physiology (Thomas et al. 2008; Zhang et al. 2008; Næsset & Gobakken 2008; Korhonen et al. 2011; Vepakomma et al. 2011). Despite these and previously cited that have investigated the potential to predict N and P with IS data in the canopy of various ecosystems including the boreal forest, we are unaware of any studies that have addressed the feasibility of remote estimation of the foliar N:P ratio. There are two reasons for this. First, the propagation of errors during arithmetic operations such as division is likely to result in low accuracy. Otherwise, the spatial representation of the N:P ratio would be a trivial task whereby a N concentration map can be divided by a P concentration map. Second,

there are no established absorption features attributable to the N:P ratio in the typical EM spectrum used for earth observation, i.e., 400-2500 nm. Therefore, any correlations between spectral data and the N:P ratio would be via a covariate that is correlated to both the remote sensing data and the N:P ratio.

In this study, we explore the potential for estimating the canopy N:P ratio for a boreal mixedwood forest using IS and LiDAR data. Our specific objectives in this study are to: 1) develop predictive models of the canopy N:P ratio using airborne and spaceborne IS data, 2) address the issues of temporal variation and the scale effect of the MAUP on prediction accuracy of the canopy N:P ratio, and 3) investigate whether LiDAR data have any explanatory power for the estimation of the canopy N:P ratio in this structurally complex mixedwood boreal forest.

2.2. Materials and Methods

2.2.1. Background

Remote sensing and field data used in this study were collected as part of the Canadian Carbon Program (formerly the Fluxnet-Canada Research Network) and some (particularly the airborne hyperspectral and LiDAR data) has been used in previous studies (e.g., Thomas et al. 2006, 2008). The motivation for this study stems from the observed significant relationships between macronutrients (N, P, magnesium, potassium) and crown closure at the same site (Gökkaya et al. 2012, in review). Since crown closure is a measure of green biomass in the canopy, remote sensing metrics such as narrowband spectral indices calculated from IS data (that have been used to predict greenness and chlorophyll) and LiDAR metrics (that have been shown to be correlated with crown closure) may have potential for modeling the N:P ratio.

2.2.2. Study Site

The study was carried out at the Groundhog River Flux Station (GRFS), located approximately 80 km southwest of Timmins, in northeastern Ontario, Canada (Figure 2.1a). A 41-m tall flux tower, which collects micrometeorological and flux data, is located at the site. This site is a mature boreal mixedwood stand generally representative of the mixedwood forest composition in the region with a heterogenous mixture of five primary tree species including trembling aspen (*Populus tremuloides* Michx.), white birch (*Betula papyrifera* Marsh.), white spruce (*Picea glauca* [Moench] Voss), black spruce, balsam fir (*Abies balsamea* [L.] Mill.), and patches of northern white cedar (*Thuja occidentalis* L.). A total of 34 plots, each 11.3 m in radius, were established using a stratified sampling scheme within the 1-km radius flux footprint of the tower to capture the major species associations. The plots included 25 circular measurement plots called Thomas plots, and nine National Forest Inventory (NFI) plots (Figure 2.1b).

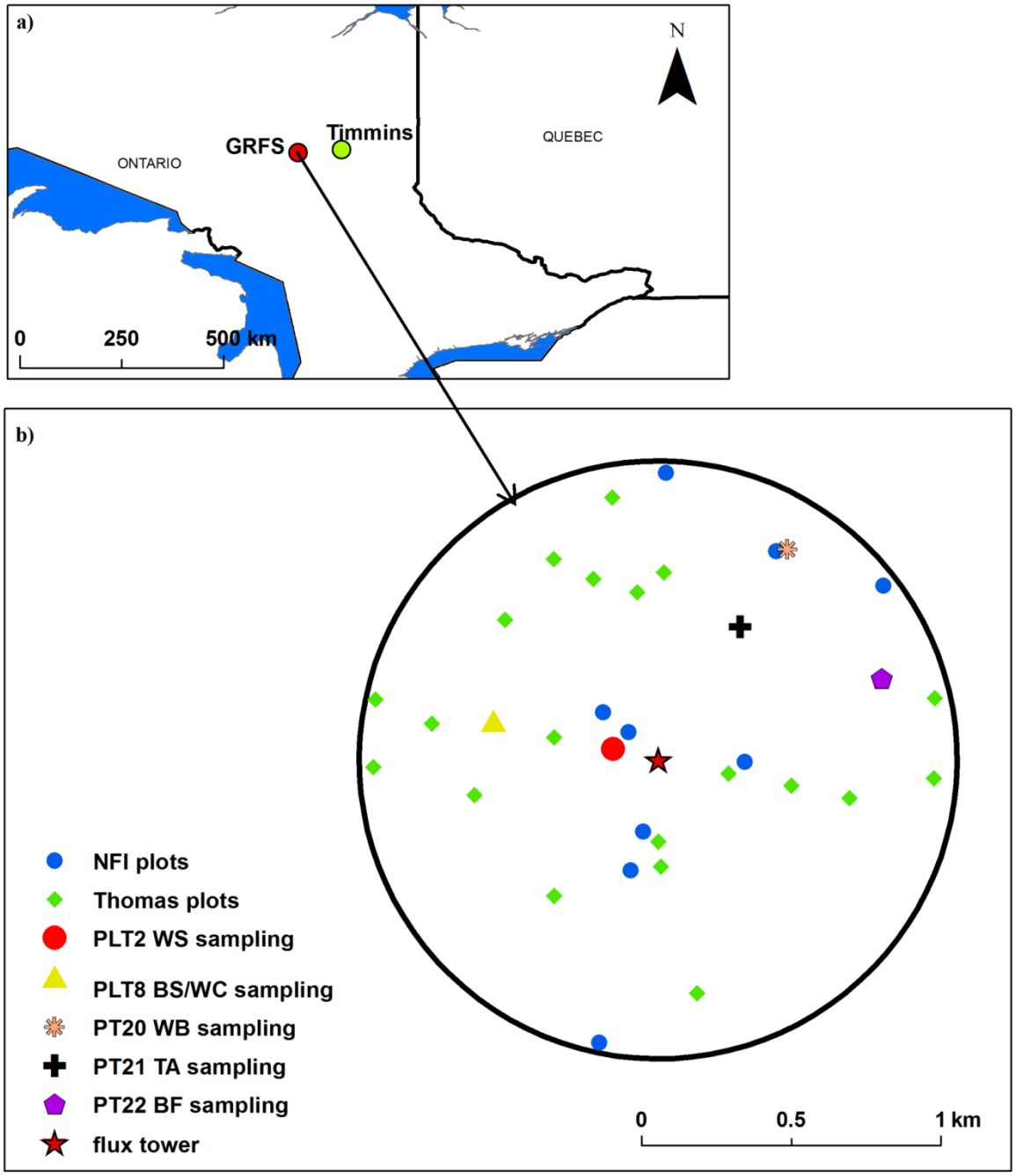


Figure 2.1. (a) Location of the GRFS in Ontario, (b) Plot layout and biochemical sampling plot locations at the mixedwood forest site. WS = white spruce (*Picea glauca*); WC = northern white cedar (*Thuja occidentalis*); BS = black spruce (*Picea mariana*); WB = white birch (*Betula papyrifera*); TA = trembling aspen (*Populus tremuloides*); BF = balsam fir (*Abies balsamea*).

2.2.3. Forest mensuration data

Forest mensuration data were collected in 2003 and 2004 in accordance network protocols (Fluxnet-Canada 2003). Diameter at breast height (dbh), height to the top of the crown, crown width, and species were recorded for each tree with a dbh greater than 9 cm within each plot. From these measurements structural metrics describing canopy shape and height including average height (arithmetic mean of all trees' heights), dominant height (maximum tree height), mean dominant height (mean height of the 100 largest trees/ha), Lorey's height (tree height weighted by basal area), basal area, and crown closure were calculated for each plot. The biomass of each species was calculated using the allometric equations developed for Ontario hardwoods and softwoods that incorporated dbh and tree height (Alemdag 1983, 1984) and reported in kg/ha for each plot.

2.2.4. Biochemical sampling and analysis

Prior work (Gökkaya et al. 2012, in review) describes the methods for foliar sample collection and analysis to determine foliar macronutrient concentration. Briefly, multiple leaf samples were collected from the sunlit portions of the canopy of five trees for each of the six species in five different Thomas plots labeled in Figure 2.1b, resulting in 30 trees being sampled in 2004 and 2005. The same trees were resampled in July 2005. Samples were analyzed at the Inorganic Laboratory of the Ontario Forest Research Institute (Sault Ste. Marie, Ontario) to determine the foliar concentration of N, P, and other macronutrients in accordance with the network protocol (Fluxnet-Canada 2003). Total N in the foliage was determined through the conversion of all forms of N into N_2 by dry combustion. The Kjeldahl method was used to extract the cations of P. Total P from the Kjeldahl extract was measured using inductively

coupled plasma-atomic emission spectroscopy. Average nutrient concentration for each species was calculated by taking the average of multiple samples collected from the five trees.

The foliar N:P ratio was calculated by dividing the average foliar N concentration by average P concentration on a dry mass basis for each species. Average N:P ratio of a given species was multiplied by its respective biomass fraction to scale the N:P ratio up from leaf to canopy level for each plot. The biomass fraction of a given species was calculated by taking the ratio of its total biomass to the biomass of all trees within each plot. Since we were more interested in the upper sunlit canopy, the use of biomass fraction instead of the number of trees of each species fraction eliminated the bias caused by many small sub-canopy trees in the plots.

2.2.5. Remote sensing data – preprocessing

Airborne hyperspectral data were collected with the Compact Airborne Spectrographic Imager (CASI) over the study site in August 2004 across a wavelength range of 400–940 nm in 72 channels with approximately 7.5 nm bandwidth. Data were converted to ground reflectance using the CAM5S atmospheric model (Freemantle 2005). Final georeferencing was completed using a sub-meter LiDAR-derived elevation model of the site, which was validated using differentially corrected GPS data (Thomas et al. 2008). Final pixel size of the georeferenced image was 4m x 4m.

Spaceborne Hyperion IS data were collected in July 2005. Hyperion EO-1 sensor collects data over 400-2500 nm spectral range with 10 nm spectral and 30 m spatial resolution, respectively. Hyperion data have lower signal-to-noise ratios (SNR) than airborne sensors; contain abnormal pixels, dropped lines, and destriping as a result of poorly calibrated or malfunctioning detectors (Datt et al. 2003) so these data require a more thorough preprocessing. First, bands that had no data or had very low SNR were removed resulting in 155 bands. Then,

surface reflectance was obtained using the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) program, which uses a radiative transfer model to convert radiance at top of atmosphere to reflectance at the surface pixel by pixel (ENVI 2009). Following atmospheric correction, the minimum noise fraction (MNF) algorithm was used to reduce the effect of noise. MNF algorithm segregates noise from the signal in the data by estimating the noise in the data. The resulting MNF bands can be divided into two parts: one part associated with large eigenvalues and spatially coherent images, which contain data, and another part with eigenvalues around one, which are noise-dominated images. By using only those MNF bands which are spatially coherent and have high eigenvalues, the noise is separated from the data, thus improving spectral processing results (Green et al. 1988; Boardman & Kruse 1994). Those MNF bands with high eigenvalues and spatially coherent images can then be used to transform back to the original spectral data space, i.e. inverse MNF transformation, sometimes after further filtering to reduce noise in the MNF bands. In this study, MNF algorithm was applied to the VNIR and SWIR bands separately because Hyperion data are acquired by different spectrometers within these two ranges and the spatial structure of the noise such as the position of destriping across the image, differs between them (Datt et al. 2003). Noise estimates were calculated from pixels collected over spectrally homogenous, clear water bodies scattered across the image. An inverse MNF transformation was carried out by using a total of 22 MNF bands based on the examination of the resulting MNF bands for eigenvalues and spatial coherence. These 22 bands were then median filtered to reduce noise prior to transformation and the VNIR and SWIR bands were combined afterwards. The image was georectified using a Landsat TM image with a total root mean squared error (RMSE) of approximately 6.5 m.

LiDAR data were collected in August 2004 by M7 Visual Intelligence Inc. using the Leica ALS-40 system (Leica Geosystems Inc., Norcross, Georgia, USA) and correlated with the upscaled canopy N:P ratios of 2004 and 2005. Since the relationship of LiDAR data to N:P ratio is through canopy structural variables, LiDAR data were used with the assumption that these structural variables do not change significantly over a one-year period barring a major disturbance event affecting this forest and therefore the LiDAR data provide a static indicator of canopy structure. The ALS-40 is a discrete return system, which captures the first and last response above a specified intensity threshold for each laser pulse (referred to as ‘‘first returns’’ and ‘‘last returns’’). This system has a pulse rate of 25 kHz and a wavelength of 1084 nm. The average point spacing is 0.7 pulses/m^2. The data were collected at an approximate flying altitude of 1067 m. The estimated horizontal and vertical positional accuracy was less than 20 cm (Thomas et al. 2006).

2.2.6. Remote sensing data – generation of hyperspectral indices and LiDAR metrics

Narrowband indices which are predictors of greenness and chlorophyll, were calculated from both the reflectance and the derivative of reflectance within the visible (400-700 nm), red-edge (680-750 nm), and the lower range of the near infrared (750-800 nm) spectra of both images. Calculating the derivative of reflectance is equivalent to the slope of the reflectance curve at any given wavelength thereby reducing the variability caused by changes in illumination or background reflectance (i.e., rock, soil, litter) (Elvidge &Chen 1995). Among these were red-edge reflectance ratio-based indices such as Gitelson and Merzlyak 2 (G&M2 = R_{750}/R_{700}), Vogelmann 2nd Index (Vog2 = $(R_{734}-R_{747})/(R_{715}+R_{726})$), greenness index (GI = R_{554}/R_{677}) and derivative indices including the maximum of derivative in the red-edge ($D_{max(680-750)}$) (R and D represent reflectance and derivative of reflectance, respectively) (Vogelmann et al. 1993;

Gitelson & Merzlyak 1997; Zarco-Tejada et al. 1999). Index values were extracted for each pixel in a plot and used to calculate a plot average. These average index values for each plot were used as predictor variables in linear regression analysis.

Raw LiDAR data are referenced to an ellipsoid or datum and need to be converted to heights above ground before relating them to forest mensuration data such as tree/canopy height. In order to obtain height above ground of the canopy, i.e. canopy height model (CHM), the ground elevation value for each LiDAR point derived from the digital elevation model (DEM) was subtracted from the individual LiDAR points. The DEM was generated using a high-density LiDAR dataset with 0.5 m resolution (Thomas et al. 2006). For each plot, LiDAR metrics that characterize the vertical and horizontal canopy structure were calculated from the first returns of the CHM. These metrics included those commonly reported for forest structure analysis, such as percentiles of height (10, 20, 30… 100), mean, standard deviation, coefficient of variation (i.e. standard deviation/mean), and the mean of the upper 25th percentile of height (e.g. Lim and Treitz 2004; Thomas et al. 2006).

2.2.7. Statistical analyses for model development

Foliar N:P ratios were compared between years using paired t-tests since the same trees were sampled in both years. Distributions of the foliar N:P ratios were tested for normality by the Shapiro-Wilk test, which is considered to be more robust than other tests for normality when the sample size is small ($n<30$) (Shapiro & Wilk 1965). Data are considered to be normally distributed when the 'W' statistic approaches 1 and 'p' is not significant. Significance threshold was set to an alpha level of 0.05.

Multiple linear regressions were conducted using the best-subsets method to generate predictive models for the N:P ratio where the indices calculated from the remote sensing data of

that year were used as predictor variables for each year. This method compares all possible models with a pre-defined number of predictor variables (Hudak et al. 2006). To avoid overfitting, we used only those models with a maximum of two variables. Multicollinearity was analyzed by calculating the variance inflation factor (VIF) with a maximum allowable VIF value of 5. Even though VIF value of 10 or greater has been proposed as an indicator of serious multicollinearity, which needs to be addressed (Kutner et al. 2004), we used a maximum VIF value of 5 as the cut-off in our analyses to be more conservative. Model explanatory power was expressed by the coefficient of determination, R^2, and the adjusted coefficient of determination, R^2_{adj}, which accounts for the improvement in explanatory power of the model as new variables are added. The residuals were tested for normality by the Shapiro-Wilk test to satisfy the normality assumption of the F distribution, which is used to test model significance. RMSE, which is an indicator of model accuracy, was reported as percent of the mean. The models were validated using the leave-one-out cross-validation method where, residuals are calculated by successively leaving one plot out of the analysis going through all the plots. The results of leave-one-out cross validation are reported by the Predicted Residual Sum of Squares (PRESS) RMSE and PRESS r^2. Low values of PRESS RMSE and high values of PRESS r^2 suggest more robust models. The significance of the variables in the model and the model itself were determined based on the p value (i.e. $p \leq 0.05$). The regression analysis was run separately with the indices and the LiDAR metrics to produce IS- and LiDAR specific models.

Lastly, we investigated whether predictive models with the same predictor variables existed both years. Having a two-year dataset allowed us to do this comparison and such models would be useful to predict how the N:P ratio values differ over time by simple model calibration.

2.2.8. Potential source of error

Two methods could have been used to predict the foliar N:P ratio via remote sensing. The first is to create modeled estimates of N and P using empirical analysis, such as those described in the statistical analysis section above. Once models of N and P are created and mapped for an area, these maps can be combined to obtain the N:P ratio across the landscape (hereafter called the map division method). This approach can potentially propagate error, resulting in lower RMSEs and significance levels of the model, since two modeled products. each with an associated error, are divided to create a third model. We examined this potential error by dividing the concentration maps of N and P generated from the index models obtained from regression analysis and comparing the coefficient of determination (r^2) and prediction errors to the upscaled foliar measurements for each plot. To avoid these propagated errors, the second, and our preferred method is to generate completely new models of the N:P ratio as described in the previous sections, which can be validated directly against the field-measured foliar N:P ratios.

2.2.9. Assessment of model performance and consideration of the MAUP

Several aspects inherent to our analysis could theoretically cause differences in model performance. The three largest factors include differences in spatial resolution (4m airborne data versus 30m satellite data), differences in radiometric quality (both the CASI and the Hyperion data suffer from noise at different wavelengths), and differences in the foliar N:P ratios from year to year. To test the robustness of indices selected for the prediction of the N:P ratio across time and space, we derived models from each dataset separately, and then applied these models to both years. This allowed for two comparisons: the low resolution indices derived from the 2005 Hyperion data were applied to the higher resolution 2004 CASI data (low → high) and the

high resolution indices derived from the 2004 data were applied to the lower resolution 2005 data (high → low).

Temporal variability of the field data fortuitously allowed us to assess the impact of the scale effect of the MAUP on model performance for predicting the canopy N:P ratio. Statistical analysis revealed the black spruce foliar N:P ratio did not vary significantly between 2004 and 2005 (Table 2.1). Given that black spruce covers the study area in large distinct homogenous patches, we were able to mask areas of pure black spruce and compare model performance for only those areas. If the model is not significantly affected by the MAUP, it should perform well to characterize the canopy N:P ratio at either resolution, with differences resulting from real differences in the foliar N:P ratio in different years.

2.3. Results

No significant differences in the value of the foliar N:P ratio were evident between years for all species grouped together. For the four coniferous species, the foliar N:P ratio did not differ significantly between 2004 and 2005. However, paper birch and trembling aspen each had significantly higher foliar N:P ratios in 2005 compared to 2004 (Table 2.1).

Table 2.1. Foliar N:P ratio values for species found at GRFS, Ontario, Canada for 2004 and 2005. Values are mean ± standard deviation. n=5 for each species. * indicates N:P ratio differs significantly between years.

	2004	**2005**
All trees	13.87 ± 2.41	14.60 ± 3.30
Populus tremuloides	14.75 ± 1.68	18.38 ± 1.81 *
Betula papyrifera	16.59 ± 2.14	18.56 ± 0.25 *
Picea glauca	12.31 ± 1.24	12.46 ± 1.28
Picea mariana	12.54 ± 0.91	13.29 ± 0.66
Abies balsamea	15.60 ± 2.14	14.72 ± 1.85
Thuja occidentalis	11.42 ± 1.71	10.17 ± 0.21

For 2004, using the high-resolution airborne data, the maximum of the derivative in the red-edge and reflectance ratio of certain wavelengths in the red-edge (Vog2 index) were shown to be the best predictors of the N:P ratio. For the 2005 satellite data, simple reflectance ratio in the red-edge (G&M 2 index) and the greenness index were identified as best for predicting the canopy N:P ratio. Both years' models were very similar in terms of variance explained (70%) for the canopy N:P ratio (Table 2.2). However, the 2004 model had a lower cross-validation PRESS RMSE error, indicating a more robust model.

Table 2.2. N:P ratio predictive model statistics generated from 2004 airborne and 2005 spaceborne IS data. The robustness of the indices as predictors of the N:P ratio across time and scale is tested by applying the high resolution indices derived from the 2004 CASI data to the coarse resolution 2005 Hyperion data (High → Low) and vice versa (Low → High). R and D represent reflectance and derivative of reflectance, respectively.

Year	Variable indices	R^2	Adj R^2	Calibration RMSE (%)	Residuals Shapiro W	Residuals Shapiro p	PRESS RMSE (%)	PRESS r^2
2004	$D_{max(680-750)}$ & $(R_{734}-R_{747})/(R_{715}+R_{726})$	0.70	0.68	4.8	0.99	0.99	5.0	0.63
2005	R_{750}/R_{700} & R_{554}/R_{677}	0.69	0.67	7	0.98	0.90	7.2	0.64
Test of robustness across time and scale								
High→ Low	$(R_{734}-R_{747})/(R_{715}+R_{726})$	0.65	0.64	7.2	0.99	0.92	7.4	0.61
Low→ High	R_{750}/R_{700}	0.34	0.32	7.0	0.97	0.55	7.3	0.24

When the indices selected for predicting the foliar N:P ratio for a particular year were tested to predict the other year's N:P ratio, only one of the indices was significant for each year and their predictive ability differed significantly. The reflectance ratio index calculated from 2004 CASI data explained 65% of the variance in the 2005 canopy N:P ratio, whereas the simple

ratio index calculated from 2005 Hyperion data explained only 34% of the variation in the 2004 canopy N:P ratio (Table 2.2).

The comparison of one and two variable predictive models between years revealed that the index of the maximum reflectance across the red-edge (R_{max}(700-770)) and modified normalized difference ratio index ((R_{750}-R_{705})/ (R_{750}+R_{705})) could explain the variance in the canopy N:P ratio with R^2_{adj} values of 0.61 and 0.51 for 2004 and 0.50 and 0.54 for 2005. Several two-variable predictive models of the canopy N:P ratio were common to both years but most failed to meet the model requirements, i.e., variable was insignificant ($p>0.05$), had high VIF, or the gain in variance explained in the canopy N:P ratio was minimal. A two-variable model common to both years had R^2_{adj} values of 0.66 and 0.60 for 2004 and 2005, respectively, for the canopy N:P ratio prediction.

The two approaches for estimating the N:P ratio, i.e. map division vs. using indices as independent variables in regression analysis, yielded different results for 2004 and 2005. The prediction accuracy increased by 68% when regression was used with high resolution 2004 data and explanatory power increased from 44 to 70%. In 2005, the prediction accuracy increase for regression vs. map division approach was only 3% with almost no change in explanatory power (Table 2.3). Lastly, the predicted canopy N:P ratio values of the pure black spruce areas were significantly greater in 2005 than 2004 (13.2 vs. 12.5, $p<0.0001$).

Table 2.3. Effect of variation in scale and estimation method (i.e. map division vs. regression) on the canopy N:P ratio prediction accuracy in 2004 and 2005.

	Map division		**Regression prediction**		**change in prediction accuracy from map division to regression (%)**
	r^2	**RMSE**	**r^2**	**RMSE**	
2004 (4m)	0.44	1.79	0.70	0.57	68
2005 (30m)	0.68	0.94	0.70	0.91	3

Although LiDAR metrics did not provide better overall predictive models for the N:P ratio over IS data, they did explain 54 and 67% of the variance in the canopy N:P ratio for 2004 and 2005, respectively. The same LiDAR metrics, i.e., 50th percentile of height and maximum height, were selected in the models for both years (Table 2.4).

Table 2.4. LiDAR predictive model statistics for the canopy N:P ratio estimation.

year	variables	R^2	Adj R^2	Calibration RMSE (%)	PRESS RMSE (%)
2004	50th percentile of height, maximum height	0.54	0.51	5.9	6.1
2005	50th percentile of height, maximum height	0.67	0.65	7.2	7.5

Crown closure is the only canopy biophysical variable that was significantly correlated with the canopy N:P ratio in both years. Other variables, including dominant height, mean dominant height, Lorey's height, basal area, and biomass exhibited significant correlations with the N:P ratio in 2005 (Table 2.5). The two height percentiles, 50th percentile of height and maximum height, calculated from LiDAR data were significantly correlated with all of the canopy biophysical variables ($p \leq 0.0001$).

Table 2.5. Correlation coefficients between canopy biophysical variables and the canopy N:P ratio. Statistical significance at 95% confidence or better is indicated by *.

	2004	2005
Dominant height (m)	0.09	0.34*
Mean dominant height (m)	0.13	0.34*
Lorey's height (m)	0.17	0.42*
Average height (m)	0.17	0.32
Crown closure	0.61*	0.73*
Basal area (m^2/ha)	0.17	0.42*
Biomass (kg/ha)	0.12	0.37*

Plot level crown closure of deciduous and coniferous species differed significantly except for cedar, which was particularly underrepresented in sample size compared to the other species. However, the crown closure within deciduous and coniferous species did not differ significantly (Table 2.6).

Table 2.6. Crown closure values of species at the plot level. Values are mean ± standard deviation. Lowercase letters denote significant differences.

species	Crown closure
Populus tremuloides	0.53 ± 0.52 a
Betula papyrifera	0.65 ± 0.45 a
Picea glauca	0.11 ± 0.09 b
Picea mariana	0.21 ± 0.18 b
Abies balsamea	0.17 ± 0.17 b
*Thuja occidentalis**	0.40 ± 0.39 ab

*Sample size varies between species based on abundance in plots. Cedar is particularly underrepresented as it was only present in two plots.

The spatial distribution of the canopy N:P ratio in 2004 and 2005 at GRFS is displayed in Figure 2.2. Both maps reveal the same pattern in that higher values of the N:P ratio corresponds to those areas dominated by deciduous species or their mixtures and lower values correspond to the areas dominated by black spruce and other conifer species.

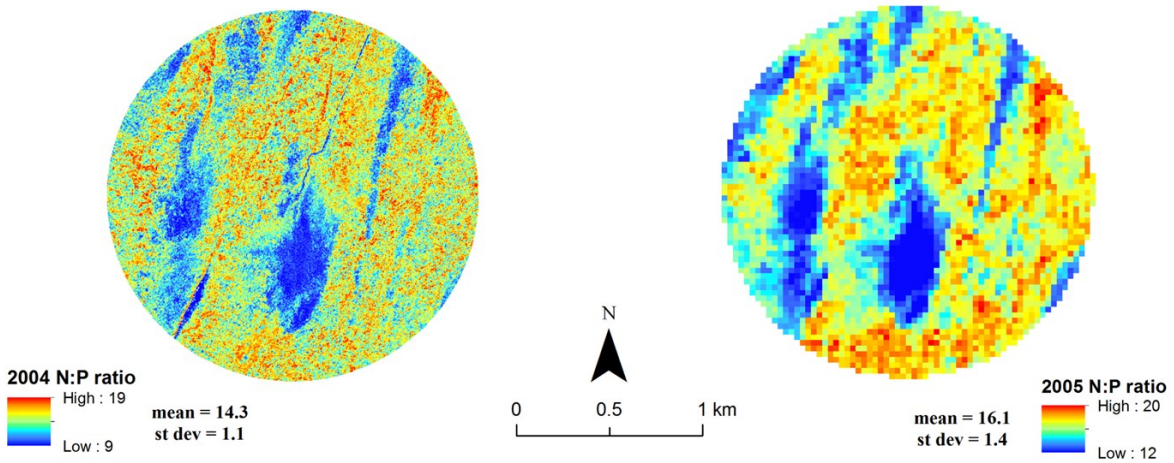


Figure 2.2. Spatial distribution of the canopy N:P ratio at GRFS in 2004 and 2005 based on the predictive models canopy N:P 2004 = 9.19 + 0.73*($D_{max(680-750)}$) – 12.59*(R_{734}-R_{747})/(R_{715}+R_{726}), and canopy N:P 2005 = 13.78 + 0.79*(R_{750}/R_{700}) – 2.14*(R_{554}/R_{677}). R and D represent reflectance and derivative of reflectance, respectively.

2.4. Discussion

2.4.1. Effect of temporal variation on prediction of the canopy N:P ratio

The main effect of the variation in the foliar N:P ratio between two years was to generate a gradient due to the differences in interspecific distribution of the foliar N:P which makes it possible to be better captured by remote sensing. The gradient in the N:P ratio among species was greater in 2005 than 2004. For example, the range for the foliar N:P ratio in 2004 was 5.2 with a coefficient of variation (CV) of 17.4% compared to 8.4 and 22.6%, respectively, for 2005. This pattern remains at the canopy level despite some reduction because of the smoothing effect of upscaling from foliar samples to plot level. Canopy level N:P ratio range and CV values were 4.6 and 8.5% in 2004 and 7.1 and 12.1% 2005. This may partially explain why the airborne data which have higher spatial and spectral resolution, produce a predictive model with almost identical explanatory power (R^2 values of 0.70 and 0.69 for 2004 and 2005 data, respectively) as the spaceborne data, which have lower spatial and spectral resolution. For the LiDAR models,

the difference is greater. The 2004 LiDAR model has an R^2 value of 0.54 whereas the 2005 model has an R^2 value of 0.67 even though the LiDAR data were collected in 2004 and would be expected to be more strongly correlated with the 2004 canopy N:P ratio. When the indices selected for one year were tested to predict the other year's data, the 2004 index originally selected for 2004 data performed better for predicting the 2005 canopy N:P ratio than vice versa, which is also probably related to the larger gradient and variability of the canopy N:P ratio in 2005. Interestingly, in 2004, balsam fir, which is coniferous, had a higher foliar N:P ratio than trembling aspen, which is deciduous but this pattern was not observed in 2005. This is possibly due to the aspen two-leaf tier moth infestation, which affected the leaf cover of the deciduous species in 2004 at this site (pers. comm. to D. Rowlinson, 2004, Provincial Forest Health Technician, Ontario Ministry of Natural Resources). Despite this disturbance, the canopy N:P ratio in 2004 could be predicted with similar accuracy to that of 2005 and shows the value of remote sensing for the N:P ratio mapping. The temporal variation in the N:P ratio should be taken into account when attempting to model it in a multitemporal setting as in this study.

2.4.2. Comparison of the two methods of predicting the canopy N:P ratio

For the higher resolution 2004 data, the two approaches for predicting the canopy N:P ratio differed significantly in explanatory power and prediction accuracy. At higher spatial resolution, variance in the N:P ratio is captured more accurately and results in much higher prediction error through error propagation and lower variance explained by the model when map division method is used. The regression method performs better with 4 m resolution data (i.e., 2004) where the prediction error decreases by 68% compared to the map division method (Table 2.3). In lower resolution 2005 data, no significant difference was evident between the two approaches. This is partly because at 30 m spatial resolution, the variance is reduced due to the

smoothing effect of larger spatial scale and as a result, error propagation is not significant for the map division method. In other words, the predictive model generated by the regression method in 2005 already had more error associated with it than the 2004 model (see Table 2.2 PRESS RMSE). Between years, the map division method generates better estimates of the canopy N:P ratio for 2005 data and the regression method did not differ in terms of variance explained but the prediction accuracy was higher for the 2004 data. It should also be kept in mind that the temporal variation was not eliminated between the 2004 and 2005 comparison. As discussed in the previous section, the gradient and variability in a given year's data may also be affecting model performance. Lastly, the radiometric quality of the two remote sensing datasets differs. Hyperion data have lower SNR than CASI data. In addition, the spectral resolution of CASI data is higher than that of Hyperion data (7.5 nm vs. 10 nm). The higher spectral resolution of CASI data allows for more refined and accurate index calculations than Hyperion data, especially for indices that report the maximum or minimum reflectance for a given range. Even though we tried to improve the radiometric quality of the data by addressing the most noticeable problems such as line drops, partial line drops and some of the striping, we could not completely ameliorate these problems inherent to the sensor characteristics.

2.4.3. Effect of the MAUP on model performance

The predicted N:P ratios of 2004 and 2005 differed significantly, ($p<0.0001$) with 2005 values being greater. This was despite the lack of significant differences between the foliar N:P ratio values of black spruce, which suggests that the scale effect of MAUP affects model performance for predicting the canopy N:P ratio. We also accounted for the imbalances in the sample size of foliar samples and the pixels and the expected effect on comparison tests. There were only five foliar samples whereas more than one hundred pixels were used to compare the

predicted canopy N:P values. Instead of only reporting significance results based on all the pixels, they were randomly grouped into samples of five and compared and this was repeated multiple times. In each case the difference between the two years was significant. The difference in spatial resolution did affect the outcome of statistical analyses as suggested by the MAUP concept. When the same area is sampled using larger areal units, the variance is reduced due to spatial aggregation.

2.4.4. Predictive models

The performance of the indices for predicting the canopy N:P ratio for the two years are similar between years regardless of data source. The R^2 values were 0.70 and 0.69 for 2004 and 2005, respectively (Table 2.2). This is encouraging for larger scale modeling attempts of the canopy N:P ratio, which can be accomplished using spaceborne IS data which covers larger areas, and is cheaper. This approach would be well-suited to the canopy N:P ratio mapping of boreal forests if it can be extended to the other types of boreal forests. Whether the estimation of the N:P ratio is possible using spaceborne IS data in different boreal forest types needs to be investigated. The existence of common indices as predictors of the canopy N:P ratio for both years is promising for predicting the canopy N:P ratio through time. Although our results are for two consecutive years on only one site, they do show that it is possible to calculate an index from the corresponding year's remote sensing data, and calibrate and use it for prediction. For both years, at least one of the indices selected in the models were simple ratio indices calculated from the wavelengths within the red-edge zone. Red-edge corresponds to the region of the EM spectrum where a sharp increase in reflectance results from a shift from pigment absorption in the red spectrum to leaf and canopy scattering in the near-infrared (Dawson & Curran 1998). The red-edge has been shown to correlate with chlorophyll, N, and leaf area index (LAI) (Curran et

al. 1990; Danson & Plummer 1995; Mutanga & Skidmore 2007). The red-edge has been found to be less sensitive to soil background and atmospheric effects and can provide information not available from a combination of near infrared and visible spectral bands (Clevers 1999), which partially explains why the red-edge indices can work for the prediction of the N:P ratio at the canopy scale. The index originally selected from the airborne sensor performed better when used to predict the N:P ratio of a different year with remote sensing data of different spatial resolution, i.e., it is temporally and spatially more robust but would need to be recalibrated when used to predict a spatially and temporally different dataset. This may be due to the improved radiometric quality and higher spectral and spatial resolution of the airborne data resulting in a more detailed spectral response of the canopy that correlates better with the canopy N:P ratio as well as the larger gradient in the N:P ratio resulting from temporal variation. These indices were originally developed to estimate chlorophyll at the leaf level using spectroradiometer data. Here, we show that they are useful in predicting the N:P ratio at the canopy level using airborne and spaceborne IS data.

Although LiDAR metrics did not provide better overall predictive models for the N:P ratio, the results provided insights into the meaning of the N:P ratio across a natural mixedwood stand where fertilization has not been applied, which is the case for much of the boreal mixedwood forest. The ecological meaning of these metrics may be related to their relationship to growth and stand productivity across species mixtures at this site. The N:P ratio is a measure of nutrient limitation of biomass production and nutrient availability and therefore directly related to growth and productivity. The variation in growth as manifested by variation in height is captured by the LiDAR metrics. The LiDAR predictive model of 2005 explains more variance in the canopy N:P ratio than that for 2004, which is probably because the N:P ratio in 2005 was

significantly correlated with canopy biophysical characteristics other than crown closure and all of these characteristics are significantly correlated with the two LiDAR metrics selected in the models (i.e., 50th percentile of height and maximum height). The second reason may again be related to the larger gradient in the canopy N:P ratio in 2005. It should also be noted that since the LiDAR data were acquired in 2004, their interpretation with regards to 2005 N:P ratio prediction results should be approached with some caution because the comparison is based on the assumption that the structure of the forest at this site does not vary significantly from the previous year.

The relationship between remote sensing data and the N:P ratio at the canopy scale is based on the correlation between a canopy structural characteristic, i.e., crown closure and the N:P ratio because unlike N or P the N:P ratio does not have absorption features in the EM. Among several field metrics, the canopy N:P ratio was significantly correlated only with crown closure for both years (Table 2.5). Crown closure varied significantly between functional groups of species at this mixedwood site (Table 2.6), and this variability was captured by both IS and LiDAR data. Crown closure is the covariate that was significantly correlated with the spectral indices and the LiDAR metrics selected in the predictive models. Whether this correlation between the N:P ratio and crown closure holds in other ecosystem types needs to be tested.

2.4.5. Spatial distribution of the canopy N:P ratio at GRFS and implications

The pattern of spatial distribution of the N:P ratio is represented in a spatially-continuous manner by the canopy N:P ratio maps (Figure 2.2). A comparison revealed that the distribution of the canopy N:P ratio at the site corresponded to the species distribution for both years. Areas dominated by deciduous species, aspen and birch, had higher N:P ratios than those dominated by the coniferous species, black spruce, white spruce, and cedar. This is expected because spruce

and cedar have lower foliar N:P ratios than the deciduous species. The 2004 map, which was produced from an airborne image with higher spatial resolution provides finer detail on spatial pattern of the N:P ratio. However, the pattern of the distribution of the canopy N:P ratio is similar.

Our results are an attempt to provide spatially-explicit maps of the canopy N:P ratio for a relatively small area, i.e., a one kilometer radius around a flux tower. Nonetheless, our results suggest that the N:P ratio map has potential diagnostic value for determining nutrient availability and limitation of biomass production for boreal mixedwoods. Future work to further assess and validate this should include a nutrient addition experiment which tests the growth response and examines the N:P ratios of vegetation. Other authors have demonstrated the value of the N:P ratio of vegetation for determining nutrient availability and limitation of biomass production in wetland and upland ecosystems (Güsewell & Koerselman 2002; Tessier & Raynal 2003; Craine et al. 2008). For example, foliar N:P ratios of less than 14 indicating N limitation and N:P ratios greater than 16 indicating P limitation of biomass production in European freshwater wetland communities were identified (Koerselman & Meuleman 1996). However, extension of these threshold values to upland systems for detecting nutrient limitation of biomass production is debatable since the species and their nutritional adaptations differ likely resulting in a different critical threshold value of the N:P ratio as indicator of nutrient limitation of biomass production. In fact, an N:P ratio of 11 was found as the critical value from the literature review of studies conducted in upland sites (Tessier & Raynal 2003). There was great variation in the N:P ratios among upland ecosystem types ranging from 7 to 29.4 based on the compilation by Tessier & Raynal (2003). Lower foliar N:P ratios with a narrower range was reported from a few studies conducted in boreal ecosystems. For example, foliar N:P ratios in pure stands of Norway spruce

and Scots pine in Sweden were 7.5 and 9, respectively and both were limited by N (Jacobson & Pettersson 2001). The foliar N:P ratio of a Norway spruce plantation in Sweden was 9.8 exhibiting colimitation by N and P (Clarholm & Rosengren-Brinck 1995) and black spruce stands in Quebec, Canada with a foliar N:P of 5 were N limited (Paquin et al. 1998). Based on these results in other boreal ecosystems and by observing the spatial pattern of the canopy N:P ratio from our maps, we suggest that the deep blue colored regions have low N availability and may be N limited since they show the lowest values of N:P ratio values. These areas are dominated by black spruce and at the GRFS, black spruce grows mostly in bog areas, which have anaerobic conditions and low pH for extended periods of the year, thus making N availability low (Bridgham et al. 1998). Based on site average N:P values of 14.3 and 16.1 in 2004 and 2005, respectively, it is not possible to determine the type of nutrient limitation of biomass production of the plant community at GRFS. Even though boreal forests are in general N limited because they have younger, less-exposed soils (Vitousek & Howarth 1991), the high variation in the critical N:P ratio values for nutrient limitation of biomass production within and across ecosystems and the scarce number of experimental studies to identify threshold values of the N:P ratio for N and P limitation of biomass production in boreal forests precludes the use of the canopy N:P ratio map for predicting nutrient limitation of biomass production at GRFS. We reiterate the need raised by Tessier & Raynal (2003) for more experimental work to identify the critical N:P ratio values for nutrient limitation of biomass production across different upland ecosystem types. It should also be noted that the use of N:P ratio threshold values for identifying nutrient limitation of biomass production of communities would be valid only when either N or P is limiting as noted by Aerts & Chapin (2000). Other factors that could limit biomass production include limitation by elements other than N or P, light availability, soil water and low

temperature. It is very unlikely that any of these factors were limiting biomass production at GRFS at the time when these samples were collected which corresponds to peak or near-peak of the growing season.

We are not aware of any other published studies that have used remote sensing data to estimate the canopy N:P ratios and the potential of remote sensing data to determine the N:P ratio needs to be tested in different ecosystems. This will require a range in N and P values across the ecosystem that is significantly greater than the variability in the regression model. This typically occurs when multiple species are present, thereby increasing the range in foliar biochemistry. In some cases, external treatments such as fertilization will generate a significant gradient which will increase the range of foliar nutrient within a given species to allow for within-species nutrient prediction and or monitoring with remote sensing data (Curran et al. 1997). At continental scales, Reich & Oleksyn (2004) reported that the foliar N:P ratio ranged from 2.6 to 111.8 over a geographical area extending from 43ºS to 70ºN latitudes based on a dataset of 1280 plant species including coniferous trees, angiosperm trees, herbs, shrubs and grasses collected from sites on six continents. The results showed that the N:P ratio was positively related to mean annual temperature and decreasing latitudes, i.e., the N:P ratio increased toward the equator and decreased toward northern latitudes. This increasing trend in the foliar N:P ratio toward the equator was observed in another study that covered a geographical range between 23.5ºS and 23.5ºN latitudes and included temperate broadleaf and coniferous and tropical forests (McGroddy et al. 2004). These authors reported N:P ratios that were almost double and with higher variation in tropical forest foliage compared to temperate forests. The higher and more variable foliar N:P ratios in tropical trees were also observed by Townsend et al. (2007). In boreal forests, the range in the N:P ratio is narrower. Foliar N:P ratios ranging from a

minimum of 5 to a maximum of 10.1 were reported from studies in pure stands of black spruce, jack pine and Norway spruce in Europe and North America (Foster & Morrison 1976; Clarholm & Rosengren-Brinck 1995; Paquin et al. 1998; Jacobson &Pettersson 2001). At the GRFS old mixedwood site, which is near 48ºN latitude, the N:P ratio ranges from 11.42 to 16.59 and 10.17 to 18.56 in two growing seasons across six boreal forest species (Table 2.1). The remote sensing estimation of the N:P ratio at the canopy level shows that the majority of the N:P values falls between 10 and 18, which is a much narrower range than reported for temperate and tropical forests. It is very likely that the remote estimation of the foliar N:P ratio is possible in temperate and tropical forest canopies since they tend to have higher N:P ratios with more variability, which is an asset for remote estimation. If remote estimation of the foliar N:P ratio is possible, then it would provide a supplementary means for diagnostic assessment of nutrient limitation of biomass production and nutrient availability and also a more comprehensive understanding of the spatial pattern of the canopy level foliar N:P ratio across wider geographical areas.

2.5. Conclusions

The possibility of modeling the canopy N:P ratio in a boreal mixedwood forest canopy by using narrowband indexes calculated from airborne and spaceborne IS was explored in this study using data from two summers (i.e., 2004 and 2005). Further, the utility of LiDAR data for modeling the canopy N:P ratio was explored. Finally, the effect of temporal variation and the difference in spatial resolution as related to the scale effect of the MAUP on the prediction accuracy of the canopy N:P ratio was examined. Our results suggest that the canopy N:P ratio can be predicted by remote sensing data based on the relationship between the canopy N:P ratio and crown closure at this site. The presence of multiple species creates a gradient in crown closure, thus making it feasible to predict the N:P ratio using remote sensing data. Moreover,

indices calculated from airborne and spaceborne IS data explain very similar variability in the canopy N:P ratio. The index originally selected from airborne IS data is temporally and spatially more robust than the index selected from the spaceborne IS data. Variation in scale resulted in significant differences in the N:P ratio prediction and temporal variation affects the prediction results based on the gradient and the variation present within a given season's N:P ratio data. Although LiDAR models did not provide improved explanatory power or prediction accuracy for predicting the canopy N:P ratio over IS data, they provide insights into the relationships of structure with growth and productivity at the site. The N:P ratio is an indicator of nutrient limitation of biomass production and nutrient availability and therefore is directly related to growth and productivity. The variation in growth as manifested by variation in height is captured by the LiDAR metrics. These results are encouraging as they show that spatially-explicit modeling of the canopy N:P ratio is possible by remote sensing data and this approach can provide a diagnostic tool for assessing nutrient limitation of biomass production and nutrient availability. Future studies should focus on experimental work to identify the critical limiting values of the N:P ratio for biomass production in boreal ecosystems as well as other upland ecosystems. The ability to predict and map the N:P ratio using remotely-sensed estimates in different wetland and upland ecosystem types will provide a better insight into the overall ecosystem response to nutrient availability and limitation.

2.6. Acknowledgements

The authors gratefully acknowledge the help of Lesley Rich, Denzil Irving, Maara Packalen, Bob Oliver, David Atkinson, Björn Prenzel, Chris Hopkinson, Laura Chasmer, Brock McLeod, Lauren MacLean, and Adam Thompson in the field. Thanks are also due to Al Cameron and Lincoln Rowlinson for establishing a cruise-line trail and for on-site assistance

with tree species identification and Garry Koteles for sharing information about tree and understory species and conducting field measurements for the National Forest Inventory (NFI) validation plots. Technical expertise for leaf sampling, processing and laboratory facilities for macronutrient analysis were provided by the Ontario Forest Research Institute (OFRI). Financial assistance for this work was provided by the Department of Forest Resources and Environmental Conservation at Virginia Tech, a USDA McIntire-Stennis Formula Grant, Queen's University in Kingston, Ontario, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canadian Carbon Program (formerly the Fluxnet-Canada Research Network) through funding from NSERC, BIOCAP Canada, and the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS).

2.7. References

Aerts, R. & Chapin, F.S. III 2000. The mineral nutrition of wild plants revisited: a re-evaluation of processes and patterns. *Advances in Ecological Research* 30: 1–67.

Arbia, G., Benedetti, R. & Espa, G. 1996: Effects of the MAUP on image classification. *Geographical Systems* 3,123–41.

Alemdag, I.S. 1983. *Mass equations and mechantability factors for Ontario softwoods*. [Petawawa National Forestry Institute Information Report PI-X-23], Canadian Forestry Service, Chalk River, Ontario, Canada.

Alemdag, I.S. 1984. *Total tree and merchantable stem biomass equations for Ontario hardwoods.* [Petawawa National Forestry Institute Information Report PI-X-46], Canadian Forestry Service, Chalk River, Ontario, Canada.

Asner, G.A. & Martin, R.E. 2008. Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment* 112: 3958-3970.

Benson, B.J. & MacKenzie, M.D. 1995. Effects of sensor spatial resolution on landscape structure parameters. *Landscape Ecology* 10: 113–120.

Boardman, J. W.& Kruse, F. A. 1994. *Automated spectral analysis: a geological example using AVIRIS data, north Grapevine Mountains, Nevada*, pp. I-407 - I-418 in Proceedings, ERIM Tenth Thematic Conference on Geologic Remote Sensing, Environmental Research Institute of Michigan, Ann Arbor, MI, US.

Bowman, W.D. 1994. Accumulation and use of nitrogen and phosphorus following fertilization in two alpine tundra communities. *Oikos* 70: 261–270.

Bridgham, S.D., Updegraff, K. & Pastor, J. 1998. Carbon, nitrogen, and phosphorus mineralization in northern wetlands. *Ecology* 79-5: 1545–1561.

Chapin, F S. 1980. The mineral nutrition of wild plants. *Annual Review of Ecology, Evolution, and Systematics* 11: 233-260.

Clarholm, M. & Rosengren-Brinck, U. 1995. Phosphorus and nitrogen fertilization of a Norway spruce forest – effects on needle concentrations and acid phosphatase activity in the humus layer. *Plant and Soil* 175: 239–249.

Clevers, J.G.P.W. 1999. The use of imaging spectrometry for agricultural applications. *ISPRS Journal of Photogrammetry and Remote Sensing* 54-5: 299–304.

Coops, N.C., Smith, M.L., Martin, M.E. & Ollinger, S.V. 2003. Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* 41: 1338-1346.

Craine, J.M., Morrow, C. & Stock., W.D. 2008. Nutrient concentration ratios and co-limitation in South African grasslands. *New Phytologist* 179: 829–836.

Curran, P. J. 1989. Remote sensing of foliar chemistry. *Remote Sensing of Environment* 30: 271−278.

Curran, P.J., Dungan, J.L. & Gholtz, H.L. 1990. Exploring the relationship between reflectance red edge and chlorophyll content of slash pine. *Tree Physiology* 7-1: 33–48.

Curran, P.J., Kupiec, J.A. & Smith, G.M. 1997. Remote sensing the biochemical composition of a slash pine canopy. *IEEE Transactions on Geoscience and Remote Sensing* 35: 415-420.

Danson, F.M. & Plummer, S.E. 1995. Red edge response to forest leaf area index. *International Journal of Remote Sensing* 16-1: 183–188.

Datt, B., McVicar, T.R., Van Niel, T.G. & Jupp, D.L.B. 2003. Preprocessing EO-1 Hyperion hyperspectral data to support the application of agricultural indexes. *IEEE Transactions on Geoscience and Remote Sensing* 41-6: 1246-1259.

Dawson, T.P. & Curran, P.J. 1998. A new technique for interpolating the reflectance red edge position. *International Journal of Remote Sensing* 19-11: 2133–2139.

Elser, J.J., Sterner, R.W., Gorokhova, E., Fagan, W.F., Markow, T.A., Cotner, J.B., Harrison, J.F., Hobbie, S.E., Odell, G.M. & Weider, L.J. 2000. Biological stoichiometry from genes to ecosystems. *Ecology Letters* 3: 540–550.

Elvidge, C.D. & Chen, Z. 1995. Comparison of broad-band and narrow-band red and near-infrared vegetation indices. *Remote Sensing of Environment* 54: 38-48.

ENVI. 2009. *Atmospheric Correction Module: QUAC and FLAASH User's Guide*. Version 4.7.

Fenn, M.E., Poth, M.A. & Johnson, D.W. 1996. Evidence for nitrogen saturation in the San Bernardino Mountains in southern California. *Forest Ecology and Management* 82: 211–230.

Fluxnet-Canada, 2003. National Forest Inventory Ground Sampling Guidelines, version 4.0. http://www.fluxnet-canada.ca/pages/protocols_en/GroundSamplingGuidelines_v.4.0._back.pdf (accessed Dec. 21, 2011).

Freemantle, J. 2005. *CASI Data Collection Report*, Sudbury 2004, Groundhog River. unpublished document, York University, Ontario, Canada.

Foster, N.W. & Morrison, I.K. 1976. Distribution and cycling of nutrients in a natural *Pinus banksiana* ecosystem. *Ecology* 57-1: 110-120.

Gitelson, A. & Merzlyak, M.N. 1997. Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing* 18: 2691-2697.

Gökkaya, K., Thomas, V., Noland, T., McCaughey, H., Morrison, I. & Treitz, P. 2012. Prediction of macronutrients using spaceborne imaging spectroscopy and LiDAR data in a mixedwood boreal forest canopy. *Canadian Journal of Remote Sensing* in review.

Gradowski, T. & Thomas, S.C. 2008**.** Responses of *Acer saccharum* canopy trees and saplings to P, K and lime additions under high N deposition. *Tree Physiology* 28-2:173-185.

Green, A.A., Berman, M., Switzer, P., & Craig, M.D. 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing* 26: 65-74.

Güsewell, S. & Koerselman, W. 2002. Variation in nitrogen and phosphorus concentrations of wetland plants. *Perspectives in Plant Ecology, Evolution and Systematics* 5: 37–61.

Han, W., Fang, J., Guo, D. & Zhang, Y. 2005. Leaf nitrogen and phosphorus stoichiometry across 753 terrestrial plant species in China. *New Phytologist* 168: 377-385.

Herbert, D.A. & Fownes, J.H. 1995. Phosphorus limitation of forest leaf area and net primary production in a highly weathered soil. *Biogeochemistry* 29: 223–235.

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P.E. & Morgan, P. 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return LiDAR and multispectral satellite data. *Canadian Journal of Remote Sensing* 32: 126–138.

Jacobson, S. & Pettersson, F. 2001. Growth responses following nitrogen and N-P-K-Mg additions to previously N-fertilized scots pine and Norway spruce stands on mineral soils in Sweden. *Canadian Journal of Forest Research* 31: 899–909.

Jelinski, D.E. & Wu, J. 1996. The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology* 11: 29–140.

Kang, H., Zhuang, H., Wu, L., Shen, G., Berg, B., Man, R. & Liu, C. 2011. Variation in leaf nitrogen and phosphorus stoichiometry in *Picea abies* across Europe: an analysis based on local observations. *Forest Ecology and Management* 261: 195–202.

Koerselman, W. & Meuleman, A.F.M. 1996. The vegetation N:P ratio: a new tool to detect the nature of nutrient limitation. *Journal of Applied Ecology* 33: 1441–1450.

Korhonen, L., Korpela, I., Heiskanen, J. & Maltamo, M. 2011. Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. *Remote Sensing of Environment* 115: 1065-1080.

Kutner, M., Nachtsheim, C., & Neter, J. 2004. *Applied Linear Regression Models*. 4th ed. McGraw-Hill.

Lim, K.S. & Treitz, P.M. 2004. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research* 19: 558–570.

Marceau, D.J., 1992. *The problem of scale and spatial aggregation in remote sensing: An empirical investigation using forestry data*. Ph.D. thesis, University of Waterloo, Waterloo, CA.

Martin, M. E. & Aber, J.D. 1997. High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological Applications* 7: 431−444.

McGroddy, M.E., Daufresne, T. & Hedin, L.O. 2004. Scaling of C: N : P stoichiometry in forests worldwide: implications of terrestrial Redfield-type ratios. *Ecology* 85: 2390–2401.

McNulty, S.G., Vose, J.M., & Swank, W.T. 1997. Scaling predicted pine forest hydrology and productivity across the southern United States. In: Quattrochi, D.A. & M.F. Goodchild, (eds.) *Scale in Remote Sensing and GIS*, pp. 187-209.

Mirik, M., Norland, J.E., Crabtree, R.L. & Biondini, M.E. 2005. Hyperspectral one-meter-resolution remote sensing in Yellowstone National Park, Wyoming: I. forage nutritional values. *Rangeland Ecology and Management* 58: 452–458.

Moody, A. & Woodcock, C.E. 1995. The influence of scale and the spatial characteristics of landscapes on land-cover mapping using remote sensing. *Landscape Ecology* 10: 363–379.

Mutanga, O. & Skidmore, A.K. 2004. Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote Sensing of Environment* 90: 104-115.

Mutanga, O., Skidmore, A.K. & Prins, H.H.T. 2004. Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. *Remote Sensing of Environment* 89: 393–408.

Mutanga, O. & Kumar, L. 2007. Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. *International Journal of Remote Sensing* 28: 4897–4911.

Mutanga, O. & Skidmore, A.K. 2007. Red edge shift and biochemical content in grass canopies. *ISPRS Journal of Photogrammetry and Remote Sensing* 62: 34–42.

Næsset, E. & Gobakken, T. 2008. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment* 112: 3079-3090.

O'Neill, R.V., Hunsaker, C.T., Timmins, S.P., Jackson, B.L., Jones, K.B., Ritters, K.H. & Wickham, J.D. 1996. Scale problems in reporting landscape patterns at the regional scale. *Landscape Ecology* 11: 169–180.

Openshaw, S. & Taylor, P.J. 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: Wrigley, N. (ed.) *Statistical applications in spatial sciences*, pp. 127-144. London, UK.

Paquin, R., Margolis, H.A. & Doucet, R. 1998. Nutrient status and growth of black spruce layers and planted seedlings in response to nutrient addition in the boreal forest of Quebec. *Canadian Journal of Forest Research* 28: 729-736.

Pax-Lenney, M. & Woodcock, C.E. 1997. The effect of spatial resolution on the ability to monitor the status of agricultural lands. *Remote Sensing of Environment* 61: 210–220.

Porder, S., Asner, G.P. & Vitousek, P.M. 2005. Ground-based and remotely sensed nutrient availability across a tropical landscape. *Proceedings of the National Academy of Sciences U.S.A.* 102: 10909–10912.

Reich, P.B. & Oleksyn, J. 2004. Global patterns of plant leaf N and P in relation to temperature and latitude. *Proceedings of the National Academy of Sciences U.S.A.* 101: 11001–11006.

Shapiro, S.S. & Wilk, M.B. 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52-3/4: 591-611.

Tessier, J.T. & Raynal, D.J. 2003. Use of nitrogen to phosphorus ratios in plant tissue as an indicator of nutrient limitation and nitrogen saturation. *Journal of Applied Ecology* 40: 523–534.

Thomas, V., Treitz, P., McCaughey, J.H. & Morrison, I. 2006. Mapping stand-level forest biophysical variables for a mixedwood boreal forest using LiDAR: an examination of scanning density. *Canadian Journal of Forest Research* 36: 34–47.

Thomas, V., Treitz, P., McCaughey, J.H., Noland, T. & Rich, L. 2008. Canopy chlorophyll concentration estimation using hyperspectral and lidar data for a boreal mixedwood forest in northern Ontario, Canada. *International Journal of Remote Sensing* 29: 1029–1052.

Townsend, A.R., Cleveland, C.C., Asner, G.P. & Bustamante, M.M.C. 2007. Controls over foliar N:P ratios in tropical rain forests. *Ecology* 88: 107–118.

Turner, D.P., Dodson, R. & Marks, D. 1996. Comparison of alternative spatial resolutions in the application of a spatially distributed biogeochemical model over complex terrain. *Ecological Modeling* 90: 53-67.

Valentine, D.W. & Allen, H.L. 1990. Foliar responses to fertilization identify nutrient limitation in loblolly pine. *Canadian Journal of Forest Research* 20: 144–151.

Vepakomma, U., St-Onge, B. & Kneeshaw, D. 2011. Response of a boreal forest to canopy opening: assessing vertical and lateral tree growth with multi-temporal lidar data. *Ecological Applications* 21: 99-121.

Vitousek, P.M. & Howarth, R.W. 1991. Nitrogen limitation on land and in the sea: how can it occur? *Biogeochemistry* 13: 87–115.

Vogelmann, J.E., Rock, B.N. & Moss, D.M. 1993. Red edge spectral measurements from sugar maple leaves. *International Journal of Remote Sensing* 14: 1563–1575.

Wessman, C.A., Aber, J.D., Peterson, D.L. & Melillo, J.M. 1988. Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. *Nature* 335: 154-156.

Zarco-Tejada, P.J., Miller, J.R., Mohammed, G.H., Noland, T.L. & Sampson, P.H.1999. "Canopy Optical Indices from Infinite Reflectance and Canopy Reflectance Models for Forest Condition Monitoring: Application to Hyperspectral CASI Data," Proceedings of the IEEE 1999 International Geoscience and Remote Sensing Symposium, June 28 – July 2, 1999, Hamburg, DE.

Zhang, Y. Chen, M.J., Miller, J.R. & Noland, T.L. 2008. Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery. *Remote Sensing of Environment* 112: 3234-3247.

Chapter 3 Prediction of macronutrients using spaceborne imaging spectroscopy and LiDAR data in a mixedwood boreal forest canopy

K. Gökkaya [a], V. Thomas [a], T. Noland [b], J.H. McCaughey [c], I. Morrison [d] and P.M. Treitz [c]

[a] Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, USA
[b] Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, ON, Canada
[c] Department of Geography, Queen's University, Kingston, ON, Canada
[d] Canadian Forest Service, Natural Resources Canada, Sault Ste. Marie, ON, Canada

This chapter is under review at Canadian Journal of Remote Sensing.

Abstract

Foliar macronutrients (nitrogen (N), phosphorus (P), potassium (K), calcium (Ca) and magnesium (Mg)) are required for plant physiological and ecosystem processes including photosynthesis (N, P, K, Mg), primary production (N), decomposition (N), respiration (N, P, K) and cell wall formation (Ca). The ability to measure, map, or model foliar macronutrients allows us to understand the spatial patterns of these processes. Imaging spectroscopy (IS) has been used for this purpose to some extent, particularly for N, using airborne and satellite imagery predominantly in temperate forest ecosystems. However, there has been very little research conducted at these scales to model P, K, Ca, and Mg and studies are lacking for boreal forests. Recent work suggests that structural information derived from light detection and ranging (LiDAR) data can also provide valuable information about foliar biochemicals at the canopy level, particularly when mapping canopy chlorophyll and carotenoids in complex multi-species environments. We report results of a study of macronutrient modeling using spaceborne IS and airborne LiDAR data for a mixedwood forest canopy composed of black and white spruce, balsam fir, northern white cedar, white birch, and trembling aspen in the boreal of northern Ontario. Predictive models using Hyperion satellite IS data were developed with adjusted R^2 of 0.73, 0.72, 0.62, 0.25, and 0.67 for N, P, K, Ca and Mg, respectively. Results for Ca suggest that IS data are not useful for predicting canopy Ca concentration in this forest type. LiDAR data are correlated with macronutrient concentrations and provide improved predictive ability over the IS models for K, Mg and particularly for Ca. The presence of multiple species at the site is driving the relationships between the LiDAR data and macronutrient concentrations by creating a gradient. The LiDAR model explained 80% of the variance in canopy Ca concentration with an

RMSE of less than 10%. When IS and LiDAR data were combined, there was an improvement for the prediction of K. Results demonstrate satellite IS data can be used for accurate modeling of the macronutrients N, P, K, and Mg at the canopy level, which is encouraging for larger scale analyses of boreal forests.

3.1. Introduction

Macronutrients (nitrogen (N), phosphorus (P), potassium (K), calcium (Ca) and magnesium (Mg)) are required for plant physiological and ecosystem processes. For example, N is an important constituent of the chlorophyll molecule and the carbon-fixing enzyme ribulose-1,5-bis-phosphate carboxylase/oxygenase and is thus directly related to photosynthesis (Field and Mooney, 1986). Foliar N is also related to primary production and decomposition (Melillo et al., 1982; Smith et al., 2002). P is a component of nucleic acids, lipid membranes, sugar phosphates and ATP which all have important roles in photosynthesis and respiration (Taiz and Zeiger, 2010). N and P are the major growth-limiting nutrients for plants worldwide (Chapin, 1980). Mg, like N, is also a constituent of the chlorophyll molecule and is directly related to photosynthesis (Taiz and Zeiger, 2010). K has a number of important roles in photosynthesis and respiration, including translocation of photosynthates into sink organs, maintenance of turgor pressure, activation of enzymes, N metabolism and reducing excess uptake of ions such as Na and Fe in saline and flooded soils (Marschner, 1995; Mengel and Kirkby, 2001). Ca is required during cell division and in the synthesis of new cell walls, particularly the middle lamellae (Taiz and Zeiger, 2010).

The ability to measure, map, or model foliar macronutrients allows us to examine these as proxies for assessing forest condition and stress, as well as ecosystem processes such as forest nutrient cycling and decomposition, carbon (C) exchange (photosynthesis and net primary production (NPP)), and to understand the spatial patterns of these processes. Photosynthesis and

NPP are of critical interest for the assessment of C capture where forests function as sinks for C in the biosphere, given that global atmospheric carbon dioxide (CO_2) concentrations have been increasing at accelerating rates over the last several decades (Olivier et al., 2011). Macronutrients, in particular N, have been used as proxies for C exchange, or as parameters in ecosystem models. For example, canopy N concentration was used to estimate the NPP of temperate forest ecosystems coupled with field measurements and hyperspectral remote sensing data by Smith et al. (2002) and as a parameter in the PnET-II forest process model (Ollinger and Smith, 2005). Thus, spatially-explicit landscape and regional inputs of foliar N concentration have the potential to improve the accuracy of ecosystem models. Maps of canopy macronutrient concentrations at small scales (e.g., in close proximity to a flux tower) can also serve as validation products for ecosystem modeling efforts.

The remote detection of forest canopy biochemistry has become feasible with the advances in remote sensing technology, in particular imaging spectroscopy (IS). The spectral response of vegetation is governed by the scattering and absorption features of internal leaf structure plus canopy structure, and biochemical constituents such as pigments, nitrogen, cellulose, lignin, proteins and water (Curran, 1989). IS provides detailed information about the spectral characteristics of objects on the Earth's surface by capturing the spectral response in numerous, narrow bands over a continuous electromagnetic spectrum (Goetz et al., 1985). Hence, it is possible to correlate absorption features related to certain biochemicals to the spectra obtained from airborne or spaceborne IS sensors using statistical methods such as regression (Curran, 1989).

Airborne IS data have been used to predict N at canopy scales predominantly in temperate forest ecosystems (Wessman et al., 1988; Curran et al., 1997; Martin and Aber, 1997).

Ollinger et al. (2008) included a boreal coniferous forest site in their study to estimate N concentration at the national scale of the US. Some studies have examined the utility of spaceborne Hyperion IS data to predict N in eucalyptus forest canopies (Coops et al., 2003) and compare the effectiveness of AVIRIS and Hyperion IS data to predict N concentration in temperate forest canopies (Smith et al., 2003; Townsend et al., 2003). However, similar studies are lacking for boreal forest ecosystems.

The detection of P, K, Mg and Ca using IS data has been more limited. The majority of studies have utilized field spectroradiometers to predict P in agricultural fields (Al-Abbas et al., 1974; Osborne et al., 2002) and P and K in the foliage of giant sequoia and eucalyptus trees (Gong et al., 2002; Ponzoni and Gonçalves, 1999). P, K, Mg and Ca concentration have been examined in savannah grasses (Mutanga et al., 2004) and willow, olive, grass and heather foliage (Ferwerda and Skidmore, 2007). At airborne scales, AVIRIS data were used to estimate area-integrated P concentrations in tropical forests of Hawaii (Porder et al., 2005) and African savannah using HyMap data (Mutanga and Kumar, 2007). Mirik et al. (2005) found a statistically significant relationship between a simple ratio vegetation reflectivity index (1129 nm/469 nm) and P concentration on an area basis in the forage vegetation of Yellowstone National Park using PROBE-1 data. The small number of studies demonstrates that there is a clear need to investigate the utility of IS data for the estimation of P, K, Mg and Ca in forest canopies. To our knowledge, there is no study that investigates the utility of spaceborne IS data for the prediction of these macronutrients. If predictive models that are comparable to the ones obtained by airborne data can be obtained by using spaceborne data, then it would be possible to predict macronutrients at the canopy scale over larger geographic areas with the launch of IS

satellites such as the HyspIRI and EnMAP with planned launch dates within this decade, which will provide continuous global coverage.

When estimations of macronutrients are conducted at canopy scale, canopy structure needs to be taken into account because the distribution of the macronutrient of interest is also related to canopy structure. Correlations between chlorophyll concentration and canopy structure (based on LiDAR metrics) have been demonstrated in mixedwood boreal forests (Thomas et al., 2008). LiDAR is an active remote sensing technique where laser pulses are usually fired off an airplane and the returns used to calculate the height of objects on the Earth's surface. For a detailed description of LiDAR remote sensing, the reader is referred to Wehr and Lohr (1999). The height information obtained from LiDAR data can be used to characterize the vertical and horizontal structure of the canopy. Metrics calculated from LiDAR data (e.g., mean, median, maximum and percentile heights, canopy density metrics, coefficient of variation) have been used extensively to characterize structural and biophysical properties of forest ecosystems such as individual and plot level tree heights, leaf area index, fractional cover and biomass; see van Leeuwen and Nieuwenhuis (2010) for a detailed review. IS and LiDAR data have been used synergistically to improve the estimate of chlorophyll and pigment concentrations in forest canopies (Blackburn, 2002; Thomas et al., 2008). As a result of its heterogeneous canopy structure, LiDAR data should prove useful in predicting canopy macronutrient concentration in boreal mixedwood forest type.

We expect to see relationships between canopy N and spectral reflectance based on its absorption features in the NIR and SWIR portions of the electromagnetic spectrum (EM) spectrum and also in the visible range because of its presence in chlorophyll. Indirect relationships between P, K, Mg and spectral reflectance are expected because of their presence in

the structure of chlorophyll and proteins that have direct absorption features in the 400-2500 nm EM spectrum range. On the other hand, Ca, which is a less mobile and a structural macronutrient, (Pallardy, 2008) may be better predicted by LiDAR data. The objectives of the study are to: 1) evaluate the utility of spaceborne IS data to estimate canopy macronutrient (N, P, K, Ca and Mg) concentrations in a mixedwood boreal forest; and 2) test the potential contribution of canopy structural information derived from LiDAR data for improving the predictive models for these macronutrients.

3.2. Materials and Methods

3.2.1. Study Site

This research was conducted at the old mixedwood site, part of the Groundhog River Flux Station (GRFS), one of the stations of the Canadian Carbon Program (formerly the Fluxnet-Canada Research Network). The site is located approximately 80 km southwest of Timmins, Ontario, Canada (Figure 3.1a) and includes a 41 m tall flux tower. The site is representative of a mature boreal mixedwood forest with a patchy mixture of five primary tree species including trembling aspen (*Populus tremuloides* Michx.), white birch (*Betula papyrifera* Marsh.), white spruce (*Picea glauca* [Moench] Voss), black spruce (*Picea mariana* [Mill.] B.S.P.), balsam fir (*Abies balsamea* [L.] Mill.), and distinct patches of northern white cedar (*Thuja occidentalis* [L.]).

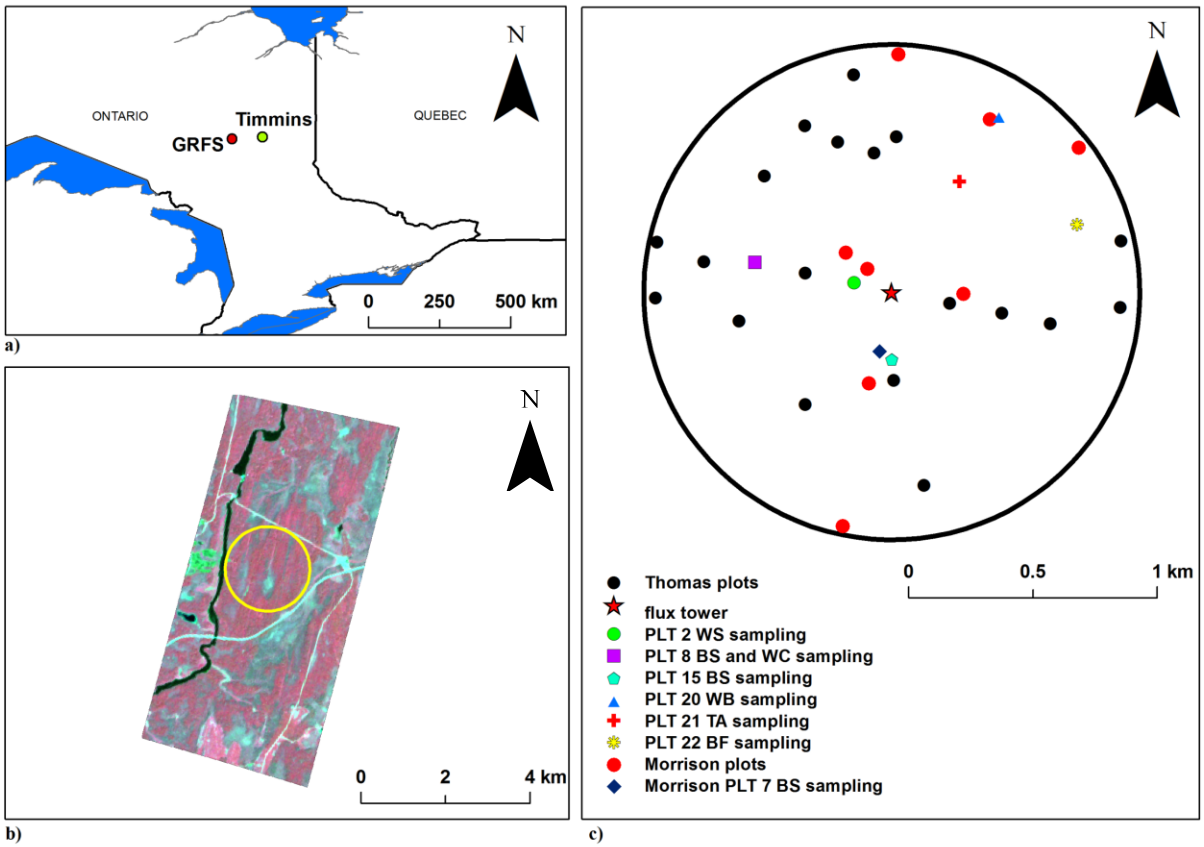


Figure 3.1. (a) Location of the GRFS, (b) Flux tower footprint at the mixedwood forest at the GRFS, (c) Plot layout and sample plot locations at the mixedwood forest site. WS = white spruce (*Picea glauca*); WC = northern white cedar (*Thuja occidentalis*); BS = black spruce (*Picea mariana*); WB = white birch (*Betula papyrifera*); TA = trembling aspen (*Populus tremuloides*); BF = balsam fir (*Abies balsamea*).

Twenty five circular measurement plots, hereafter called Thomas plots, 11.3 m radius, were established using a stratified sampling scheme within the 1 km radius flux footprint of the tower to capture the major species associations. Similarly, nine National Forest Inventory (NFI) plots, hereafter called Morrison plots, were established (Figure 3.1b, c).

3.2.2. Field data

Multiple leaf samples were collected from the upper sunlit portions of the canopy from five trees in several plots for all species except black spruce for which 13 trees were sampled in July 2005 (Figure 3.1c). The samples were analyzed at the inorganic laboratory of the Ontario

Forest Research Institute (Sault Ste. Marie, Ontario) to determine the foliar concentration of N, P, K, Ca and Mg. Total N in the foliage was determined through the conversion of all forms of N into N_2 by dry combustion. The Kjeldahl method was used to extract the cations of P, K, Ca and Mg, which were measured using inductively coupled plasma - atomic emission spectroscopy. Further details of these methods can be found in the Ontario Forest Research Institute Laboratory Standard Operating Procedures reports 112 and 113 (OMNR, 2007a,b). Average nutrient concentration for each species was calculated by averaging the results from the five tree samples for all species except black spruce for which 13 trees were sampled. To scale the macronutrient concentration up from leaf level to canopy level for each plot, the following formula was used:

$$\overline{chl}_{plot} = \sum_{i=1}^{n} (\overline{chl}_i) f_i$$

Where i = species in plot, $\overline{chl}_i$= the mean total chlorophyll concentration for species i, and f_i = biomass fraction of species i within the plot. We eliminated the bias caused by many small sub-canopy stems in the plot by using the biomass fraction, as opposed to the number of stems of each species fraction since we were more interested in the upper sunlit canopy. Forest mensuration data were collected in 2003 and 2004 in accordance with Fluxnet-Canada protocols (Fluxnet-Canada, 2003; now known as the Canadian Carbon Program). Diameter at breast height (dbh), height to the top of the crown, crown width, and species were recorded for each tree with a dbh greater than 9 cm within each plot. Structural metrics describing canopy shape and height were calculated from these measurements including average height (arithmetic mean of all trees' heights), dominant height (maximum tree height), mean dominant height (the mean height of the 100 largest trees/ha), Lorey's height (tree height weighted by basal area), basal area, and crown closure were calculated for each plot. Total aboveground biomass for each plot was calculated

summing the aboveground biomass of trees in the plot using the allometric equations developed for Ontario hardwoods and softwoods that incorporated dbh and tree height (Alemdag, 1983; 1984) and was reported in kg/ha.

3.2.3. Remote sensing data

Hyperion EO-1 data (400-2500 nm with 10 nm spectral resolution and 30m spatial resolution) were collected for the GRFS in July 2005. Hyperion data have a lower signal-to-noise ratio (SNR) than airborne sensors, hence requiring significant preprocessing. First, bands that had no data or had very low SNR were omitted from analysis. Second, scan lines that caused image striping were restored by taking the average of the adjacent lines. Subsequently, the data were atmospherically corrected using the fast line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) program. FLAASH uses the MODTRAN4 radiation transfer code to convert radiance at the top of the atmosphere to reflectance at the surface, pixel by pixel (ENVI, 2009). Following atmospheric correction, geometric correction was completed using a high-density (3-8 pulses m^{-2}) discrete return LiDAR dataset collected at a flying altitude of 244 m in August, 2003 by Airborne 1 Corp. (El Segundo, CA, USA) using the ALTM 2050 (Optech Inc., Toronto, Ontario, Canada) with positional error of 13 cm in the x-y direction and less than 15 cm in the z direction (Optech Inc, 2002). The final root mean squared error (RMSE) of the georectification was less than one third of a Hyperion pixel size, i.e. less than 10m. The same LiDAR data were used to calculate the metrics that were used in the regression analysis.

3.2.4. Model development

Hyperion reflectance data and LiDAR metrics that aid in characterizing canopy structure were calculated for each plot.

When developing models for N, initially the absorption features in the NIR (750-1400 nm) and SWIR (1400-2500 nm) associated with N-H bonds located at wavelengths 1020, 1510, 1980, 2060, 2130, 2180, 2300 nm, were used in multiple linear regression analysis to avoid bands that have spurious correlations (Curran, 1989). In the second approach, the visible wavebands (400-700 nm), which contain chlorophyll absorption features and the red-edge portion of the spectrum (680-750 nm) that is correlated with chlorophyll were included in the analysis (Horler et al., 1983; Filella and Peñuelas, 1994).

The following spectral regions/features were used for P model development: (i) absorption features across the 550-750 nm and 2015-2195 nm ranges (Mutanga and Kumar, 2007); (ii) the simple ratio vegetation index (1129 nm/462 nm) (Mirik et al., 2005); and (iii) the wavelengths available from Hyperion data that are adjacent to these wavelengths. To develop predictive models for Mg, wavebands in the visible portion of the spectrum that contain chlorophyll absorption features and the red-edge were used. This approach was taken because Mg, like N, is a component of chlorophyll and therefore correlations with spectral information would likely be because of the presence of chlorophyll. For K and Ca, wavebands within the whole range of the Hyperion spectral coverage were used.

3.2.5. Relationships between macronutrients and canopy structure

Canopy macronutrients and canopy structure relationships were examined at the plot level to help identify the useful LiDAR metrics. The LiDAR metrics were calculated from the first returns of the canopy height model for each plot. These metrics included commonly reported metrics for analysis of forest structure, such as percentiles of height (10, 20, 30… 100), coefficient of variation, mean, and the mean of the upper 25th percentile height (e.g., Lim and Treitz, 2004; Thomas et al., 2008).

3.2.6. Regression Analysis

Multiple linear regression was conducted using the best subsets approach to generate predictive models for the macronutrients where IS reflectance data and LiDAR metrics were used as predictor variables. This approach compares all the possible models with a pre-defined number of predictor variables (Hudak et al., 2006). To avoid overfitting, models with a maximum of two variables were developed. Multicollinearity was avoided by allowing a maximum variance inflation factor (VIF) of 5. Model explanatory power was expressed by the coefficient of determination, R^2, and the adjusted coefficient of determination, R^2adj, which accounts for the improvement in explanatory power of the model as new variables are added. The residuals were tested for normality using the Shapiro-Wilk test to satisfy the normality assumption of the F distribution, which is used to test for model significance. RMSE, an indicator of model accuracy, was reported as percent of the mean. Model validation was performed using the leave-one-out cross-validation method where residuals are calculated by successively leaving one plot out of the analysis for validation purpose (Holiday et al., 1995). The results of leave-one-out cross validation are reported by the predicted residual sum of squares (PRESS), and PRESS RMSE statistics. Low values of these statistics suggest more robust models. The significance of the variables in the model and the model itself was tested by the p value where the probability, p, was less than 0.05. The regression analysis was applied separately to IS and LiDAR data to produce models that contained only IS and only LiDAR metrics. Finally, IS reflectance data and LiDAR metrics were used together in the regression analysis to test whether there was any improvement from the combined approach.

3.3. Results

Foliar macronutrient concentrations of the principal species on the site are shown in Table 3.1. Trembling aspen and white birch tend to have the highest concentrations of macronutrients, whereas black and white spruce have the lowest concentrations. In all cases, Ca is the exception as it exhibits a random distribution of concentration across species.

Table 3.1. Site and species average macronutrient concentrations (mean ± standard deviation) for the mixedwood site at the Groundhog River Flux Station, 2005 (n=5 for all species except *Picea mariana* whose n=13).

	N (%)	P (ppm)	K (ppm)	Ca (ppm)	Mg (ppm)
All trees	1.23 ± 0.64	873 ± 264	4647 ± 1764	8175 ± 2275	1435 ± 668
Populus tremuloides	2.16 ± 0.09	1184 ± 115	7070 ± 1444	11970 ± 1990	2556 ± 362
Betula papyrifera	2.28 ± 0.19	1229 ± 110	5547 ± 734	5885 ± 1380	2222 ± 473
Abies balsamea	1.24 ± 0.09	857 ± 154	5096 ± 963	9828 ± 998	1151 ± 302
Picea glauca	0.94 ± 0.15	755 ± 103	3259 ± 415	6569 ± 1022	892 ± 232
Picea mariana	0.65 ± 0.11	612 ± 106	3074 ± 825	7643 ± 1506	1010 ± 254
Thuja occidentalis	1.04 ± 0.05	1023 ± 50	6355 ± 563	8007 ± 1007	1459 ± 197

The macronutrient predictive models based on IS data show high R^2s for all nutrients except Ca (Table 3.2a). There is a slight improvement for N prediction when the visible and red-edge wavebands are included with the spectral channels specific to the N absorption features. P is predicted most accurately among all macronutrients (i.e., possessing the lowest RMSE and a high R^2 value). The Mg and K predictive models have adjusted R^2 values of 0.67 and 0.62 with RMSEs of 15.4 % and 14.5%, respectively.

Table 3.2. Predictive models generated from: (a) IS reflectance data, (b) LiDAR data, and (c) IS and LiDAR data combined.

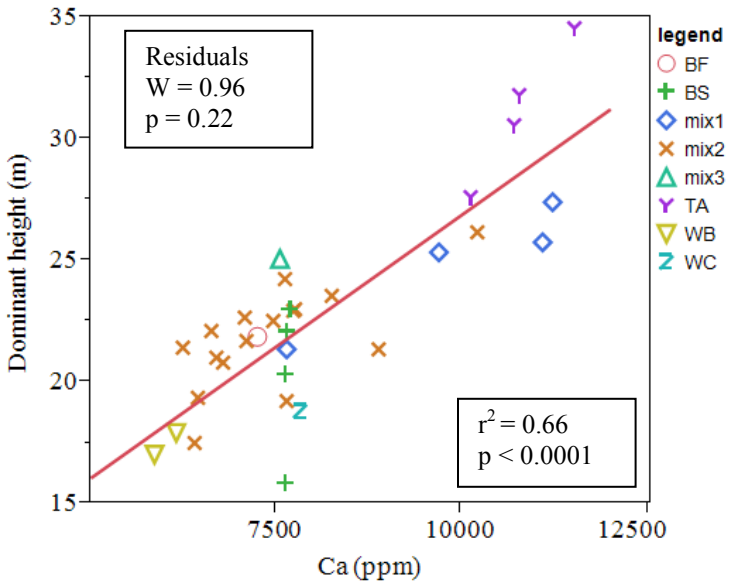
Nutrient	Variables	R^2	Adjusted R^2	F ratio	p	Calibration RMSE (%)	Residuals Shapiro W	Residuals Shapiro p	PRESS RMSE (%)
a) IS reflectance data									
	Wavelengths (nm)								
N	1023, 2173	0.70	0.68	35.5	<0.0001	19	0.98	0.69	22.9
N	590, 1507	0.75	0.73	46.2	<0.0001	17.3	0.96	0.32	18.1
P	548, 1134	0.74	0.72	44.4	<0.0001	10.6	0.96	0.32	11.9
K	548, 1679	0.64	0.62	27.4	<0.0001	14.5	0.96	0.28	15.2
Ca	457, 2214	0.29	0.25	6.4	0.0046	17.1	0.98	0.62	18.1
Mg	548, 752	0.69	0.67	35.2	<0.0001	15.4	0.98	0.62	16.9
b) LiDAR data									
	LiDAR metric								
N	50th perc. ht., max. ht.	0.71	0.69	37.6	<0.0001	18.7	0.97	0.37	19.5
P	50th perc. ht., max. ht.	0.65	0.63	29	<0.0001	12.3	0.96	0.30	12.9
K	50th perc. ht.	0.68	0.67	67.7	<0.0001	13.5	0.97	0.45	13.9
Ca	50th perc. ht., covar.	0.80	0.79	61.7	<0.0001	9.1	0.98	0.75	9.4
Mg	50th perc. ht., max. ht.	0.77	0.76	53	<0.0001	13.3	0.96	0.32	13.8
c) IS and LiDAR data									
N	1023, 50th perc. ht.	0.75	0.74	47.7	<0.0001	17.1	0.96	0.26	18.5
P	IS model above	0.74	0.72	44.4	<0.0001	10.6	0.96	0.32	11.9
K	1679, 50th perc. ht.	0.76	0.74	47.8	<0.0001	12	0.94	0.05	12.7
Ca	LiDAR model above	0.80	0.79	61.7	<0.0001	9.1	0.98	0.75	9.4
Mg	752, 50th perc. ht.	0.78	0.77	55.7	<0.0001	13	0.98	0.91	14

Strong correlations between Ca and dominant height, as well as N, P, K, Mg and crown closure ($p<0.0001$) (Table 3.3) led to further exploration as to the predictive power of these field metrics for canopy macronutrients. A highly significant ($p<0.0001$) relationship between Ca and dominant height (Figure 3.2a) and N and crown closure was found (Figure 3.2b). When plots were grouped by species dominance, different patterns emerged for the two graphs. Plots dominated by trembling aspen and the mixture of trembling aspen and white birch represented the maximum values of dominant height and Ca concentration and white birch dominated plots represented the lowest values of both. Plots dominated by black spruce and the mixtures of deciduous and coniferous species had intermediate values (Figure 3.2a). For crown closure and N concentration relationship, blacks spruce and mixed conifer plots had the lowest values. White birch dominated plots in addition to trembling aspen and trembling aspen and white birch dominated plots represented the highest values. Intermediate values were represented by deciduous and coniferous species dominated plots (Figure 3.2b). These canopy structural metrics were significantly correlated to the LiDAR metrics, in particular the 50th percentile of height and maximum height (Table 3.4).

Table 3.3. Spearman's correlation coefficients between macronutrient concentration and canopy structural metrics (* and ** denote significance at $p<0.01$ and $p<0.0001$, respectively).

	N (%)	**P (ppm)**	**K (ppm)**	**Ca (ppm)**	**Mg (ppm)**
Dominant Height (m)	0.34	0.33	0.48**	0.71*	0.47**
Mean Dominant Height (m)	0.34	0.33	0.49**	0.70*	0.46**
Lorey's Height (m)	0.42	0.41	0.52**	0.65*	0.51**
Average Height (m)	0.32	0.30	0.41	0.48**	0.39
Crown Closure	0.74*	0.74*	0.69*	0.18	0.76*
Basal Area (m^2/ha)	0.45**	0.46**	0.63*	0.56**	0.55**
Biomass (kg/ha)	0.38	0.38	0.52**	0.48**	0.48**

a)



b)

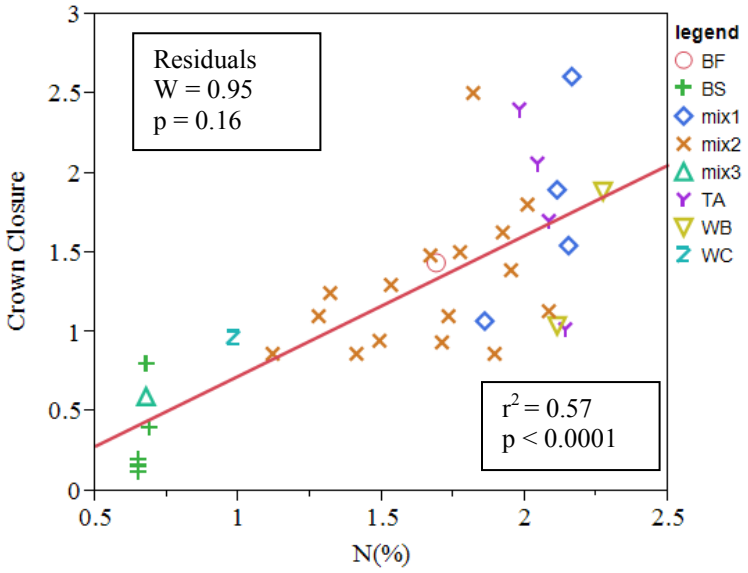


Figure 3.2. The relationship between: (a) Ca concentration and dominant height, and (b) N concentration and crown closure. W stands for the Shapiro-Wilk test statistic. BF=balsam fir, BS=black spruce, mix1=mixed deciduous, mix2=mixed deciduous and conifer, mix3=mixed conifer, TA=trembling aspen, WB=white birch, WC=white cedar.

Table 3.4. Spearman's correlation coefficients between canopy structural metrics and LiDAR metrics selected for the predictive models (* and ** denote significance of p<0.01 and p<0.0001, respectively).

	50th percentile of height	**Maximum height**	**Coefficient of variation**
Dominant Height (m)	0.74*	0.92*	-0.28
Mean Dominant Height (m)	0.81*	0.92*	-0.34
Lorey's Height (m)	0.82*	0.90*	-0.34
Average Height (m)	0.76*	0.80*	-0.39
Crown Closure	0.65*	0.43**	-0.68*
Basal Area (m^2/ha)	0.78*	0.61*	-0.57**
Biomass (kg/ha)	0.81*	0.74*	-0.46**

The predictive models generated from the LiDAR metrics are summarized in Table 3.2b. The LiDAR metrics provided the most significant improvement for the prediction of Ca compared to the IS models (adjusted R^2 = 0.79; RMSE = 9.1%). The adjusted R^2 for the Ca predictive model increased more than three-fold and the RMSE decreased 50%. The second noticeable improvement was for the prediction of K, where the adjusted R^2 of the model increased to 0.67 from 0.62 despite being a univariate model, and the RMSE decreased (Table 3.2b). A univariate model was selected for K because the second variable in the bivariate model was either not significant or, the increase in the adjusted R^2 of the bivariate model was minimal. Finally, the prediction of Mg using LiDAR metrics also improved compared to the IS models. The adjusted R^2 increased to 0.76 from 0.67 and the RMSE decreased. For N, there was not a marked difference compared to the IS models. P is not predicted as well with the LiDAR metrics with the adjusted R^2 decreasing 12.5% compared to the IS data model.

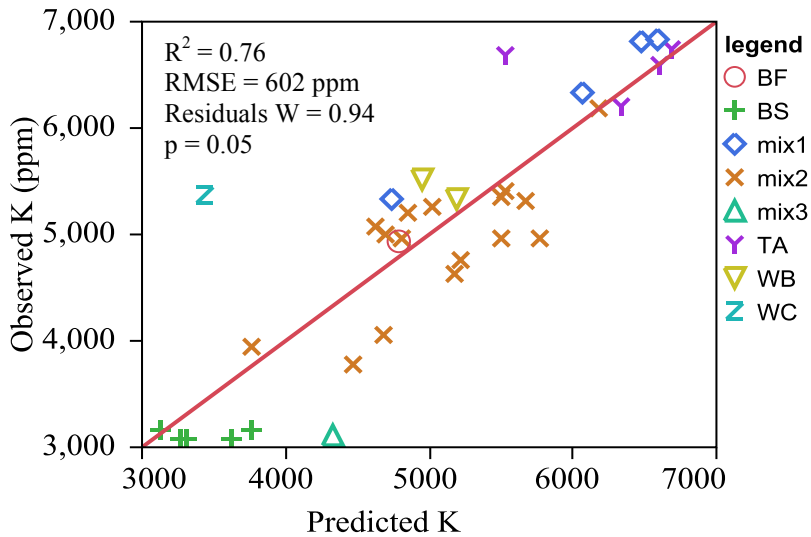
The only noticeable improvement over the IS- or LiDAR-only models was for the prediction of K when both the IS and LiDAR data were used (Table 3.2c). The adjusted R^2 increased 19% and 10% over the IS- and LiDAR-only models, respectively. The calibration RMSEs decreased by 17% and 11% and the cross validation RMSEs decreased by 16% and 9%

compared to the IS- and LiDAR-only models, respectively. Predicted values of K concentration show good agreement with observed K concentrations. Plots dominated by trembling aspen and mixtures of trembling aspen and white birch represented the highest values of K concentration and black spruce dominated plots represented the lowest values of K concentration. White birch and mixture of deciduous and coniferous dominated plots constituted the intermediate values of K concentration (Figure 3.3a). The spatial distribution of canopy K concentration tends to mimic the species distribution at the site with areas dominated by trembling aspen and white birch exhibiting higher canopy K concentrations, and areas dominated by black spruce exhibiting lower concentrations (Figure 3.3b). The average K concentration for the site is 4686 ppm with a standard deviation of 982 ppm. The IS model generated the best P prediction and the LiDAR model was optimal for the prediction of Ca. Combined data models produced no noticeable improvement for the prediction of Mg or N over the IS- or LiDAR-only models.

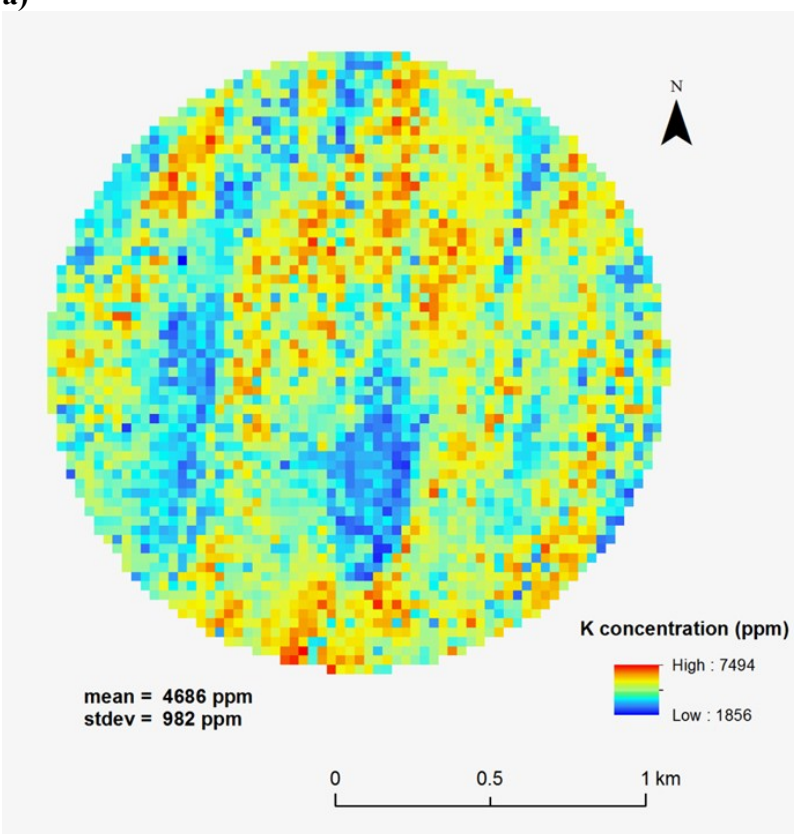
There were also strong correlations among the macronutient concentrations with the exception of Ca which did not exhibit significant correlations with N, P and K concentrations (Table 3.5).

Table 3.5. Spearman's correlation coefficients among the macronutrient concentrations of foliar samples collected at the Groundhog River Flux Site in July 2005 (* and ** denote significance at $p<0.05$ and $p<0.0001$, respectively).

	P	K	Ca	Mg
P	1.00			
K	0.81*	1.00		
Ca	0.15	0.30	1.00	
Mg	0.79*	0.68*	0.38**	1.00
N	0.85*	0.70*	0.16	0.63*



a)



b)

Figure 3.3. (a) Observed versus predicted K concentration obtained from the IS and LiDAR model in Table 2c with the equation K = -606 + 247*$R_{1679\ nm}$ + 162*50th perc. ht. W stands for the Shapiro-Wilk test statistic. BF=balsam fir, BS=black spruce, mix1=mixed deciduous, mix2=mixed deciduous and conifer, mix3=mixed conifer, TA=trembling aspen, WB=white birch, WC=white cedar. (b) Spatial distribution of canopy K concentration at the mixedwood forest site calculated from this model.

3.4. Discussion

3.4.1. IS predictive models

The visible range of the spectrum has predictive value for N because it shows the N-containing chlorophyll absorption features. For P prediction model, one of the variables selected was 548 nm band corresponding to the green peak reflectance, which has been linked to chlorophyll and N (Gitelson and Merzlyak, 1996; Peñuelas et al., 1994). Another important variable was the 1134 nm band in the NIR region, which has been correlated with P concentration (Mirik et al., 2005) (Table 2a). Similarly, Mg model contains wavelengths that have been related to chlorophyll prediction (548 nm; 752 nm) (Table 2a). The green peak band (548 nm) and the 1679 nm wavelength in the SWIR region were selected for the K model (Table 2a). Green reflectance is related to chlorophyll and the 1690 nm band attributed to lignin, starch, protein and N is the absorption feature nearest the 1679 nm band. The canopy K concentration might be related to this band via proteins and its role in activation of coenzymes. As reported by Mutanga et al. (2004) for African savanna grasses, the macronutrients exhibit strong correlations among themselves (with the exception of Ca) (Table 5). This partially explains why the 548 nm wavelength was useful for the P, K and Mg models. No absorption features for Ca, K, Mg or P have been established in the visible, NIR and SWIR portions of the spectrum. The correlations we observed are likely a result of indirect relationships between these macronutrients and the chemicals that have known absorption features including chlorophyll, N and proteins.

Despite the low SNR characteristic of Hyperion data compared to airborne sensor data, our estimates of macronutrients are comparable to the few other studies that were conducted using spectroradiometer and airborne data. Airborne IS data produced an R^2 of 0.63 and RMSE of 28% for canopy P concentration in African savannah grass (Mutanga and Kumar, 2007).

These authors found that including the 2015-2195 nm SWIR range decreased their RMSE to 28% from 58%. Our estimates for P were not improved when the SWIR region wavelengths were used. Differences between the statistical methods used in the two studies (i.e., regression versus neural network analysis), the lower SNR of Hyperion data within the SWIR range and the difference between the ecosystem types may explain this difference. The simple ratio vegetation index (1129 nm/462 nm) calculated from airborne IS data was correlated with P ($R^2 = 0.65$) in the forage vegetation canopy composed of grasses, sedges, forbs, sagebrush and willow in Yellowstone National Park (Mirik et al., 2005). This index plus the individual wavelengths were included in our regression analysis of P. The spectral band at 1134 nm (wavelength closest to 1129 nm available from Hyperion data) was picked as the second variable in the model. Predictive models with R^2 values of 0.73, 0.67, 0.77, 0.17 and 0.33 for N, Ca, Mg, P and K, respectively, were obtained for greenhouse-grown tropical grass canopies using continuum removal technique in the 550-750 nm range using spectroradiometer data (Mutanga et al., 2005). Our results yielded better predictive models for P and K, but not for Ca (Table 2a). One reason for this difference might be related to the low correlations of Ca with other macronutrients in our dataset (Table 5) compared to theirs where Ca exhibited very significant correlations with N and Mg. Likewise, the correlations of N with K and P in their study were weaker than in our dataset. Since the 550-750 nm range contains pigment absorption features which are indirectly related to N, the lower predictive accuracies for K and P and better predictive accuracy for Ca might be attributed to the strength of correlation with N. Visible, NIR and SWIR portions of the spectrum predicted N, Ca, Mg, P and K with R^2 values of 0.70, 0.50, 0.68, 0.80 and 0.64, respectively, in the canopy of five grass species using spectroradiometer data and continuum removal technique (Mutanga et al., 2004). Our results are comparable with the exception of Ca. Another study

estimated N, Ca, Mg, P and K with R^2 values of 0.76, 0.65, 0.56, 0.85 and 0.73, respectively, from spectroradiometric reflectance collected from olive, heather, willow and mopane leaf samples (Ferwerda and Skidmore, 2007). These models were developed using stepwise regression where up to four bands were included in the models. In the studies cited sofar, the prediction accuracy for Ca is consistently better than our findings. In addition to the proposed explanations as to the possible differences, one other important factor may be related the differences in functional types, i.e., grass versus tree. Since Ca is a structural element and also stored in large quantities in woody tissue (Pallardy, 2008), the received signal from the canopy layer may not contain enough information to generate good predictive models for Ca in this forest type. This is supported by the lack of significant relationship between Ca concentration and crown closure. Future studies should investigate the effect of functional type differences for the estimation of Ca.

The ability to model N, P, K and Mg using spaceborne IS data in this mixedwood boreal forest demonstrates the potential for mapping these macronutrients at the canopy scale across larger geographic areas with the availability of data from IS satellites that are planned to be launched within this decade, which will provide global IS data. This could be a useful approach for the boreal forests of Canada which cover broad geographical ranges if predictive models may be generated in other types of boreal forests across its distribution range. Future studies that address this possibility should be conducted.

Multiple linear regression has been criticized because of the possibility of overfitting during the calibration process and selection of bands with spurious correlations. Firstly, for N we addressed this by focusing on those wavebands centered on or closest to the absorption features that have been attributed to N. Secondly, in the case of P and Mg, we focused on the portions of

the electromagnetic spectrum that contain absorption features indirectly related to the macronutrient based on the previous reports. Lastly, we limited the number of predictor wavelengths to a maximum of two.

3.4.2. Relationships between macronutrients and canopy structure and LiDAR models

The strong relationships observed between macronutrients and LiDAR metrics is based on the significant correlations between canopy structure (and how well the LiDAR point cloud distribution samples that structure) and macronutrient concentration. For example, Ca concentration could be predicted with an r^2 value of 0.66 by dominant height and N concentration could be predicted by crown closure with an r^2 value of 0.57 (Figures 3.2a and b). The relationship between N, P, K and Mg and crown closure suggests that the amount of green leaf material in the canopy may be the underlying driver of this relationship. Significant correlations exist between chlorophyll and these macronutrients with the exception of Ca (not shown). This paves the way for future research examining greenness and chlorophyll indices for the prediction of N, P, K and Mg.

The correlations between the concentration of K, Mg and Ca and canopy structural variables are reflected in the improvement of their predictive models generated from LiDAR metrics compared to the IS models (Table 3.2b). The best improvement in prediction accuracy achieved by using LiDAR metrics was in Ca probably because of its significant correlation with dominant height (Figure 2a). LiDAR metrics did not improve the prediction of N and P compared to the IS models, likely because they lacked significant correlations with height measurements. In fact, LiDAR metrics decreased the prediction accuracy for P (Table 3.2b). Previously, strong correlations were observed between chlorophyll and carotenoid concentrations and height measurements and crown closure at this site (Thomas et al., 2008). The vertical and

horizontal canopy heterogeneity of the GRFS boreal mixedwood forest, which is a result of the presence of multiple species, makes it possible to capture canopy structure by the LiDAR data and relate it to macronutrient concentration. Plots tend to be separated based on their dominant species when graphed to examine the relationship between canopy structural variables and nutrient concentration (Fig 3.2 a,b). The presence of multiple species is creating the gradient required for predicting macronutrients. Whether this relationship extends to different forest types and other ecosystems remains to be determined and requires further study.

When IS data and LiDAR metrics were used in combination to predict macronutrient concentrations, the IS model proved optimal for predicting P whereas the LiDAR model was optimal for predicting Ca. At the same time, there was no improvement observed for predicting N and Mg. K was the only macronutrient whose predictive accuracy increased with the combination of IS and LiDAR data (Table 3.2c). Since the best prediction of Ca requires LiDAR data, this suggests that Ca is strongly related to canopy structure, at least for this forest type.

It is worthwhile to note that the LiDAR data were collected two years before foliar sample collection but we think the time separation will not affect the results because the relationship of LiDAR data to macronutrients are through canopy structure variables such as crown closure and canopy height. We expect these field measured variables to not differ significantly from year to year unless there is a big disturbance event between the years. We are not aware of any big disturbance on this forest during this period.

3.5. Conclusions

This study has tested the utility of spaceborne IS and airborne LiDAR data for modeling canopy macronutrient concentrations in a boreal mixedwood forest canopy. In doing so, we have evaluated the contribution of canopy structural information provided by LiDAR data to predict

the concentrations of the macronutrients. It was possible to model canopy N, P, K and Mg concentrations using spaceborne Hyperion data and demonstrates the potential for mapping these macronutrients at the canopy scale across larger geographic areas especially with the availability of data from IS satellites that are planned to be launched within this decade, which will provide global IS data. LiDAR data are related to macronutrient concentrations through relationships to canopy structure, specifically canopy height and crown closure. The presence of multiple species at the site is driving the relationships between the LiDAR data and macronutrient concentrations by creating a gradient. LiDAR data improved the prediction of Ca significantly at this study site. Finally, the predictive ability for K is improved when the IS and LiDAR data are combined.

3.6. Acknowledgements

The authors gratefully acknowledge the help of Lesley Rich, Denzil Irving, Maara Packalen, Bob Oliver, David Atkinson, Björn Prenzel, Chris Hopkinson, Laura Chasmer, Brock McLeod, Lauren MacLean, and Adam Thompson in the field. Thanks are also due to Al Cameron and Lincoln Rowlinson for establishing a cruise-line trail and for on-site assistance with tree species identification and Garry Koteles for sharing information regarding tree and understory species and conducting all of the field measurements for the National Forest Inventory (NFI) validation plots. Technical expertise for leaf sampling, processing and laboratory facilities for macronutrient analysis was provided by the Ontario Forest Research Institute (OFRI). Financial assistance for this work was provided by the Department of Forest Resources and Environmental Conservation at Virginia Tech, a USDA McIntire-Stennis Formula Grant, Queen's University in Kingston, Ontario, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canadian Carbon Program (formerly the Fluxnet-Canada

Research Network) through funding from NSERC, BIOCAP Canada, and the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS).

3.7. References

Al-Abbas, A.H., Barr, R., Hall, J.D., Crane, F.L., and Baumgardner, M.F. 1974. Spectra of normal and nutrient deficient maize leaves. *Agronomy Journal*, Vol. 66, pp. 16–20.

Alemdag, I.S. 1983. *Mass equations and mechantability factors for Ontario softwoods*. Canadian Forestry Service, Petawawa National Forestry Institute, Chalk River, Ontario. Information Report PI-X-23.

Alemdag, I.S. 1984. *Total tree and merchantable stem biomass equations for Ontario hardwoods*. Canadian Forestry Service, Petawawa National Forestry Institute, Chalk River, Ontario. Information Report PI-X-46.

Blackburn, G.A. 2002. Remote sensing of forest pigments using airborne imaging spectrometer and LIDAR imagery. *Remote Sensing of Environment*, Vol. 82, pp. 311–321.

Chapin, F S. 1980. The mineral nutrition of wild plants. *Annual Review of Ecology and Systematics*, Vol. 11, pp. 233-260.

Coops, N.C., Smith, M.L., Martin, M.E., and Ollinger, S.V. 2003. Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 6, pp. 1338-1346.

Curran, P. J. 1989. Remote sensing of foliar chemistry. *Remote Sensing of Environment*, Vol. 30, pp. 271−278.

Curran, P.J., Kupiec, J.A., and Smith, G.M. 1997. Remote sensing the biochemical composition of a slash pine canopy. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 35, No. 2, pp. 415-420.

ENVI. 2009. *Atmospheric Correction Module: QUAC and FLAASH User's Guide*. Version 4.7

Ferwerda, J.G., and Skidmore, A.K. 2007. Can nutrient status of four woody plant species be predicted using field spectrometry? *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 62, pp. 406-414.

Field, C., and Mooney, H.A.1986. The photosynthesis–nitrogen relationship in wild plants. In *On the Economy of Plant Form and Function: Proceedings of the Sixth Maria Moors Cabot Symposium, Evolutionary Constraints on Primary Productivity, Adaptive Patterns of Energy Capture in Plants*, August 1983, Harvard Forest, USA. Edited by T.J. Givnish. Cambridge University Press, Cambridge, pp. 25–55.

Filella, I., and Peñuelas, J. 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *International Journal of Remote Sensing*, Vol. 15, No. 7, pp. 1459-1470.

Fluxnet-Canada, 2003. *National Forest Inventory Ground Sampling Guidelines*, version 4.0. http://www.fluxnet-canada.http://www.fluxnet-canada.ca/pages/protocols_en/GroundSamplingGuidelines_v.4.0._back.pdf (accessed Oct. 31, 2011).

Gitelson, A., and Merzlyak, M.N. 1996. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *Journal of Plant Physiology*, Vol. 148, pp. 501–508.

Goetz, A.F.H., Vane, G., Solomon, J.E., and Rock, B.N.1985. Imaging spectrometry for earth remote sensing. *Science*, Vol. 228, No. 4704, pp. 1147-1153.

Gong, P., Pu, R., and Heald, R.C. 2002. Analysis of in situ hyperspectral data for nutrient estimation of giant sequoia. *International Journal of Remote Sensing*, Vol. 23, No. 9, pp. 1827–1850.

Holiday, D.B., Ballard, J.E. and McKeown, B.C. 1995. PRESS-related statistics: regression tools for cross-validation and case diagnostics. *Medicine and Science in Sports and Exercise*, Vol. 27, pp. 612-620.

Horler, D.N.H., Dockray, M., Barber, J., and Barringer, A.R. 1983. Red edge measurements for remotely sensing plant chlorophyll content. *Advances in Space Research*, Vol..3, No. 2, pp. 273–277.

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P.E., and Morgan, P., 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return LiDAR and multispectral satellite data. *Canadian Journal of Remote Sensing*, Vol. 32, No. 2, pp. 126–138.

Lim, K.S., and Treitz, P.M. 2004. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research*, Vol. 19, No. 6, pp. 558–570.

Marschner, H. 1995. *Mineral Nutrition of Higher Plants*. 2nd Edition, Academic Press, San Diego.

Martin, M. E., and Aber, J.D. 1997. High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological Applications*, Vol. 7, pp. 431−444.

McCaughey, J.H., Pejam, M.R., Arain, M.A., and Cameron, D.A. 2006. Carbon dioxide and energy fluxes from a boreal mixedwood forest ecosystem in Ontario, Canada. *Agricultural and Forest Meteorology*, Vol. 140, pp. 79–96.

Melillo, J. M., Aber, J.D., and Muratore, J.M. 1982. Nitrogen and lignin control of hardwood leaf litter decomposition dynamics. *Ecology*, Vol. 63, pp. 621–626.

Mengel, K., and Kirkby, E. A. 2001. *Principles of Plant Nutrition*. 5th Edition, Kluwer Academic Publishers, Dordrecht.

Mirik, M., Norland, J.E., Crabtree, R.L., and Biondini, M.E. 2005. Hyperspectral one-meter-resolution remote Sensing in Yellowstone National Park, Wyoming: I. forage nutritional values. *Rangeland Ecology and Management*, Vol. 58, pp. 452–458.

Mutanga, O., Skidmore, A.K., and Prins, H.H.T. 2004. Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. *Remote Sensing of Environment*, Vol. 89, pp. 393–408.

Mutanga, O., Skidmore, A.K., Kumar, L., and Ferwerda, J. 2005. Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain. *International Journal of Remote Sensing*, Vol. 26, No. 6, pp. 1093–1108.

Mutanga, O., and Kumar, L. 2007. Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. *International Journal of Remote Sensing*, Vol. 28, No. 21, pp. 4897–4911.

Olivier, J.G.J., Janssens-Maenhout, G., Peters, J.A.H.W., and Wilson, J. 2011. *Long-term trend in global CO_2 emissions 2011 report*. PBL/JRC, The Hague.

Ollinger, S.V., and Smith, M.L. 2005. Net primary production and canopy nitrogen in a temperate forest landscape: An analysis using imaging spectroscopy, modeling, and field data. *Ecosystems*, Vol. 8, pp. 1–19.

Ollinger, S.V., A. D. Richardson. M. E. Martin, D. Y. Hollinger, S. E. Frolking, P.B.Reich, G. G. Katul, L. C. Plourde , J. W. Munger, R. Oren, M.L. Smith, K. T. Paw U, P. V. Bolstad, B. D. Cook, M. C. Day, T. A. Martin, R. K. Monson, and Schmid, H.P. 2008. Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proceedings of the National Academy of Science*s, Vol. 105, No. 49, pp. 19336-19341.

OMNR., 2007a. *Simultaneous analysis of total carbon and total nitrogen in foliage by dry combustion and thermal conductivity detection method*. Ontario Ministry of Natural Resources, OFRILS SOP: 112.

OMNR., 2007b. *Analysis of major cations in foliage by modified Kjeldahl digestion and inductively coupled plasma emission spectrometer detection method*. Ontario Ministry of Natural Resources, OFRILS SOP: 113.

Optech, Inc. 2002. *ALTM 2050 Airborne Laser Terrain Mapper*. On-line technical specifications document, available at http://www.ntsinfo.com/inventory/images/ALTM_2050Optech.Ref702.pdf (accessed Oct 31, 2011).

Osborne, S.L., Schepers, J.S., Francis, D.D., and Schlemmer, M.R. 2002. Detection of phosphorous and nitrogen deficiencies in corn using spectral radiance measurements. *Agronomy Journal*, Vol. 94, pp. 1215–1221.

Pallardy, S.G. 2008. *Physiology of Woody Plants*. 3rd Edition, Academic Press, San Diego.

Peñuelas, J., Gamon, J.A., Fredeen, A.L., Merino, J., and Field, C.B. 1994. Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sensing of Environment*, Vol. 48, pp. 135–146.

Ponzoni, F.J., and Gonçalves, J.L. de M.1999. Spectral features associated with nitrogen, phosphorus, and potassium deficiencies in *Eucalyptus saligna* seedling leaves. . *International Journal of Remote Sensing*, Vol. 20, No. 11, pp. 2249-2264.

Porder, S., Asner, G.P., and Vitousek, P.M. 2005. Ground-based and remotely sensed nutrient availability across a tropical landscape. *Proceedings of the National Academy of Science*s, Vol. 102, No. 31, pp. 10909–10912.

Smith, M.L., Ollinger, S.V., Martin, M.E., Aber, J.D., Hallett, R.A., and Goodale, C.L. 2002. Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. *Ecological Applications*, Vol. 12, pp. 1286−1302.

Smith, M.L., Martin, M.E., Ollinger, S.V., and Plourde, L. 2003. Analysis of hyperspectral data for estimation of temperate forest canopy nitrogen concentration: Comparison between an airborne (AVIRIS) and a spaceborne (Hyperion) sensor. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 6, pp. 1332−1337.

Taiz, L., and Zeiger, E. 2010. *Plant Physiology*. 5th Edition, Sinauer Associates Inc., Sunderland.

Thomas, V., Treitz, P., McCaughey, J.H., Noland, T., and Rich, L. 2008. Canopy chlorophyll concentration estimation using hyperspectral and lidar data for a boreal mixedwood forest in northern Ontario, Canada. *International Journal of Remote Sensing*, Vol. 29, pp. 1029–1052.

Townsend, P.A., Foster, J.R., Chastain, Jr., R.A., and Currie, W.S. 2003. Imaging spectroscopy and canopy nitrogen: Application to the forests of the central Appalachian Mountains using Hyperion and AVIRIS. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 6, pp. 1347−1354.

van Leeuwen, M., and Nieuwenhuis, M. 2010. Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research*, Vol. 129, No.4, pp. 749–770.

Wehr, A., and Lohr, U. 1999. Airborne laser scanning - an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 54, pp. 68–82.

Wessman, C.A., Aber, J.D., Peterson, D.L., and Melillo, J.M. 1988. Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. *Nature*, Vol. 335, pp. 154-156.

Chapter 4 Testing the robustness of predictive models for chlorophyll generated from spaceborne imaging spectroscopy data in mixedwood boreal forest canopy

K. Gökkaya [a], V. Thomas [a], T. Noland [b], J.H. McCaughey [c], I. Morrison [d] and P.M. Treitz [c]

[a] Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, USA
[b] Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, ON, Canada
[c] Department of Geography, Queen's University, Kingston, ON, Canada
[d] Canadian Forest Service, Natural Resources Canada, Sault Ste. Marie, ON, Canada

Abstract

The amount of chlorophyll in the leaf influences photosynthetic potential and can be an indicator of the overall condition of a plant, including its stress level and nutritional status. Therefore, understanding the spatial and temporal variation in chlorophyll concentration is important. Imaging spectroscopy (IS) made it possible to estimate chlorophyll both at leaf and canopy levels. Spaceborne imaging spectrometers became available in the last decade and provide the possibility of estimating chlorophyll concentration at larger spatial scales at less cost. We undertook this study to test the robustness of predictive models generated using Hyperion data for predicting chlorophyll concentration of datasets rom different locations collected in different years. Two indices, the derivative chlorophyll index (DCI = D_{705}/D_{722}) and the maximum derivative of the red-edge band divided by the derivative of 703 nm ($D_{max(680-750)})/D_{703}$), emerged as good predictors of chlorophyll concentration across time and space. When the canopy level predictive models of these two indices with the same parameters as generated from Hyperion data were applied to other years' data for chlorophyll prediction, they explained 71, 63 and 6% of the variation in chlorophyll for DCI and 61, 54 and 8 % of the variation in chlorophyll for $D_{max(680-750)})/D_{703}$ in 2002, 2004 and 2008, respectively with prediction errors ranging from 11.7% to 14.6%. Two-variable models generated using 2005 Hyperion data were not as robust for predicting chlorophyll concentration from other years. Only two of the five best models could explain more than half of the variance in 2004 chlorophyll concentration. Single and two variable models applied to 2008 chlorophyll data provided poor predictions. The reason is the high degree of overlap in chlorophyll concentration distribution among species. The presence of multiple species creates a gradient in the chlorophyll concentration and this gradient makes it

possible to predict chlorophyll concentration as well as affecting the performance of predictive models generated using data from a different year. Results suggest that predictive models obtained from Hyperion data are robust in predicting chlorophyll concentration within the same site across time and sensors and across space as well.

4.1. Introduction

Chlorophylls are the most important pigments. The amount of chlorophyll in plant leaves influences photosynthetic potential and primary production (Demming-Adams and Adams, 1996). Chlorophyll also contains a large proportion of total leaf nitrogen, in its molecular structure, second only to the carbon-fixing enzyme ribulose-1,5-bis-phosphate carboxylase/oxygenase (Rubisco). Foliar chlorophyll concentration generally decreases under stress (Hendry et al., 1987; Gitelson and Merzlyak, 1996). Therefore, foliar chlorophyll concentration is an indicator of plant condition including stress levels and nutritional status.

Given the importance of foliar chlorophyll for plant development and growth, understanding its spatial and temporal variation has received attention. Although wet laboratory techniques for quantifying chlorophyll exist, they are time and labour intensive and extrapolation from a limited number of leaf samples to canopy level is likely to introduce inaccuracies (Blackburn, 2007). In this context, remote sensing provides an easier, cheaper and more accurate alternative for estimating chlorophyll at canopy level. Advances in remote sensing technology, in particular, imaging spectroscopy has made remote detection of pigment concentration feasible. The physical basis for this is the existence of absorption features of pigments in the reflectance spectra of leaves (Curran, 1989; Ustin, 2009). Chlorophylls have absorption peaks in the blue and red regions of the spectrum. The blue peak is not used for estimating chlorophyll because it overlaps with absorbance of the carotenoids and xanthophylls. Chlorophyll absorbance features

are centered at the red wavelengths of 640 and 660 nm. The red-edge region (680-750 nm), a region of transition from low reflectance due to chlorophyll absorption in the red to high reflectance due to cell structure in the near infrared (NIR) is also highly correlated with chlorophyll (Horler et al., 1983; Curran et al., 1990; Filella and Peñuelas, 1994; Blackburn, 1999; Thomas et al., 2008). Since imaging spectroscopy provides detailed spectral information by capturing the spectral response in numerous, narrow bands over a continuous electromagnetic spectrum, it is sensitive to the specific absorption features of chlorophyll.

A common method of applying IS data to estimate chlorophyll is the use of spectral indices calculated from spectroradiometer, airborne and spaceborne platforms (Gitelson and Merzlyak, 1996; Datt, 1998; Sims and Gamon, 2002; Coops et al., 2003; Thomas et al. 2008; Wu et al., 2010). A comprehensive review of indices developed prior to 2003 to predict chlorophyll was compiled by le Maire et al. (2004). Most indices use ratios of narrow bands within spectral ranges that are sensitive to chlorophylls to those bands that are chlorophyll insensitive and/or are related to some other controls on reflectance. This approach solves the problems of overlapping absorption spectra of different pigments, leaf surface interactions, and variation in leaf and canopy structure (Chappelle et al., 1992; Peñuelas et al., 1995).

The availability of high spectral resolution data acquired from airborne platforms made it possible to estimate chlorophyll concentration in forest canopies (Curran et al., 1997; Zarco-Tejada et al., 2004; Asner and Martin, 2008; Thomas et al., 2008), agricultural crops (Wu et al., 2010) and a Mediterranean ecosystem (Stagakis et al., 2010). Characterization of canopy chlorophyll distribution at various scales allows processes such as gross primary production (GPP) to be quantified in agricultural and boreal forest ecosystems (Gitelson et al., 2006; Wu et al., 2009; Zhang et al. 2009).

One of the most common statistical methods used in predicting chlorophyll concentration both at leaf and canopy scales has been regression (Filella and Peñuelas, 1994; Curran et al., 1997; Coops et al., 2003; Thomas et al., 2008; Stagakis et al., 2010; Wu et al., 2010). However, a major drawback of this empirical approach is that these models don't provide accurate results when used for another location or sensor, or are applied at the same location over time (Curran, 1994; Gobron et al., 1997; Asner et al., 2003). In this study, we attempted to address this problem by testing how well predictive models derived from satellite remote sensing data predicted chlorophyll concentration for boreal mixedwood forest stands. By doing so, we attempted to 1) identify spectral indices that can be used predict canopy chlorophyll concentration across years, 2) test the robustness of the models using these indices to predict canopy chlorophyll concentration across years within the same site, and 3) extend the applicability of these predictive models of canopy chlorophyll concentration to a different geographic location and time.

4.2. Materials and Methods

4.2.1. Study sites

This research was conducted at two sites: the Groundhog River Fluxnet Site (GRFS) located approximately 80 km southwest of Timmins, in Ontario, Canada and the Sudbury area in Ontario, Canada (Figure 4.1). The GRFS hosts one of the stations of the Fluxnet-Canada Research Network. It is representative of a mature boreal mixedwood forest with a heterogenous mixture of five primary tree species: trembling aspen (*Populus tremuloides* Michx.), white birch (*Betula papyrifera* Marsh.), white spruce (*Picea glauca* [Moench] Voss), black spruce (*Picea mariana* [Mill.] B.S.P.), balsam fir (*Abies balsamea* [L.] Mill.) and distinct patches of northern white cedar (*Thuja occidentalis* [L.]). At the GRFS site, 34 plots were established within the 1

km radius flux tower footprint using a stratified sampling scheme. 11 plots were established representing pure stands of four boreal tree species: jack pine (*Pinus baksiana* Lamb.), black spruce, trembling aspen and white birch were established north of Sudbury, located approximately 200 km south of Timmins, Ontario. All were circular plots 20 m in diameter.

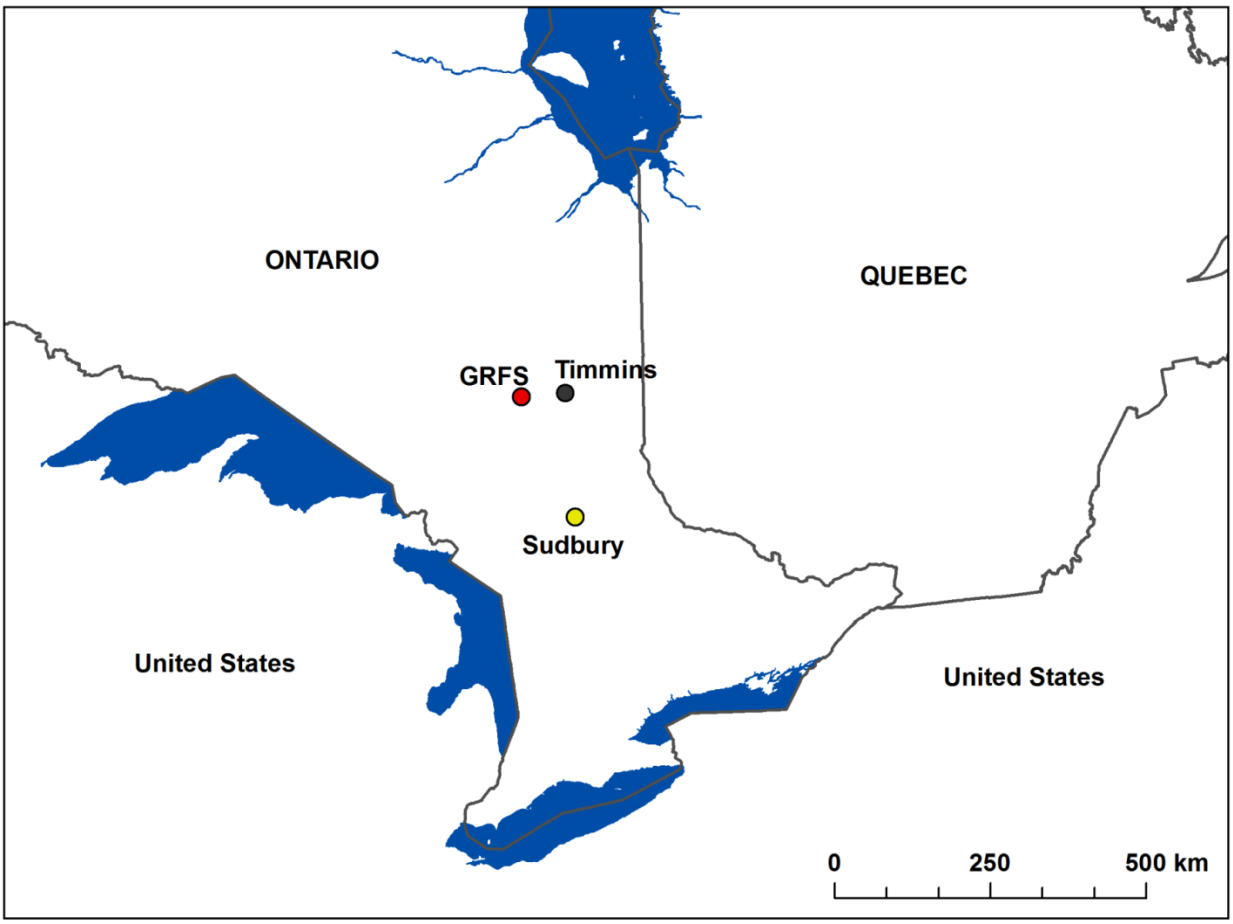


Figure 4.1. Location of the Groundhog River Fluxnet Site (GRFS) and Sudbury in Ontario, Canada. Plots were established within the 1-km footprint of the flux tower at GRFS and just north of Sudbury.

4.2.2. Biochemical sampling and analysis

Chlorophyll concentration of leaves was determined using leaf samples collected from each of the study sites. At GRFS, multiple leaf samples were collected by shotgun in July 2005 from the sunlit portions of the canopy of multiple trees for each of the six species: in total 40

trees were sampled for chlorophyll analyses. The selected trees were located in and around plots with relatively homogenous species composition, with the exception of white spruce which was sampled from a mixed species plot because the site lacked homogenous patches of white spruce (Figure 4.1). The same criteria were used to collect additional foliar samples in July 2008. Balsam fir was not sampled in 2008 because it is a sub-canopy species at this site. Sudbury area plots were sampled in July 2002 from pure stands of the four boreal species. Five trees per species were sampled in each plot resulting in a total of 55 samples. In all cases, foliage was frozen in coolers with dry ice and samples remained frozen until chlorophyll analysis was performed at the biochemistry laboratory of the Ontario Forest Research Institute (Sault Ste. Marie, Ontario) using standard laboratory techniques and spectrophotometric analyses as outlined by Wellburn (1994). Mean total chlorophyll concentration ($\overline{chl}$) was determined for each species. Plot averages for canopy chlorophyll concentration were then calculated as follows:

$$\overline{chl}_{plot} = \sum_{i=1}^{n} (\overline{chl}_i) f_i$$

Where: i = species in plot, $\overline{chl}_i$= the mean total chlorophyll concentration for species i, and f_i = biomass fraction of species i within the plot. By using the biomass fraction, as opposed to the number of stems of each species fraction, we eliminated the bias caused by many small sub-canopy stems in the plot. This is important because we are more interested in the upper sunlit canopy. For the GRFS plots, tree biomass was calculated using the allometric equations developed for Ontario hardwoods and softwoods that incorporated diameter at breast height and tree height (Alemdag, 1983; 1984). The biomass fraction for each species in each plot was determined by dividing the total plot-level species biomass by the total plot biomass. Since the

Sudbury area plots were located in pure stands (i.e., $f = 1$), the $\overline{chl}$ was used. In addition, chlorophyll data collected from the GRFS site in 2004 (Thomas et al., 2008) was used in identifying common indices for predicting chlorophyll across years and testing model robustness.

4.2.3. Remote sensing data

Spaceborne Hyperion data were collected in July 2002 (Sudbury) and 2005 (GRFS) to coincide with foliar sample collection in the field. The Hyperion EO-1 sensor collects data over the 400 to 2500 nm spectral range with 10 nm spectral and 30 m spatial resolution, respectively. Hyperion data have lower signal-to-noise (SNR) ratio than airborne sensors; contains abnormal pixels, dropped lines, and destriping due to poorly calibrated or malfunctioning detectors (Datt et al., 2003) so it requires a more through preprocessing. First, bands without data or with very low SNR were removed resulting in 155 bands. Then, surface reflectance was obtained using the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) program, which uses a radiative transfer model to convert radiance at top of atmosphere to reflectance at the surface pixel by pixel (ENVI, 2009). Following atmospheric correction, the minimum noise fraction (MNF) algorithm was used to reduce the effect of noise. This algorithm segregates the noise from the signal by estimating the noise in the data. The resulting MNF bands were divided into two parts: one part consisting of large eigenvalues and spatially coherent images, which contain data, and a second part with eigenvalues around one, which are noise-dominated images. By using only those MNF bands with high eigenvalues, the noise is separated from the data, thus improving spectral processing results (Green et al. 1988; Boardman and Kruse, 1994). Those MNF bands with high eigenvalues and spatially coherent images can then be transformed back to the original spectral data space, i.e., inverse MNF transformation, sometimes after further filtering to reduce noise in the MNF bands. Because of differences in the spatial structure of

noise due to the position of destriping as well as Hyperion data being acquired by different spectrometers within the visible and near-infrared (VNIR) and shortwave infrared (SWIR) regions, we applied the MNF algorithm separately to the VNIR and SWIR bands (Datt et al., 2003). Noise estimates were calculated from pixels collected over spectrally homogenous, clear water bodies scattered across the image. An inverse MNF transformation was carried out using a total of 22 MNF bands based on examination of the resulting MNF bands for eigenvalues and spatial coherence. These 22 bands were then median filtered to reduce noise prior to transformation and the VNIR and SWIR bands were combined afterwards. The 2002 and 2005 Hyperion images were georectified using the corresponding year's Landsat TM image as the base image with total root mean squared errors (RMSE) of approximately 2 and 6.5 m, respectively. In July 2008, airborne AVIRIS data were acquired for the GRFS area. Data were converted to surface reflectance and georectified at the University of New Hampshire Complex Systems Research Center. The spectral and spatial resolution of AVIRIS data are 10 nm and 17 m, respectively.

4.2.4. Generation of hyperspectral indices

Narrowband indices that are predictors of greenness and chlorophyll were calculated from both the reflectance and the derivative of reflectance within the visible (400-700 nm), red-edge (680-750 nm), and the lower range of the near infrared (750-800 nm) spectra of 2002 Hyperion, 2005 Hyperion and 2008 AVIRIS images. A subset of these indices is listed in Table 4.1. The same indices calculated from airborne IS data for chlorophyll prediction in 2004 were used for to identify common indices to use for predicting chlorophyll across years and testing the robustness of the models generated from 2005 data. See Thomas et al. (2008) for a detailed description of these data. In the current study, 34 plots (including the extra plots used for

validation by Thomas et al. 2008) instead of 24 were used for the regression analysis and no transformations were applied to the data. The derivative of reflectance is equivalent to the slope of the reflectance curve at any given wavelength which reduces the variability due to changes in illumination or background reflectance (i.e., rock, soil, litter) (Elvidge and Chen, 1995). Index values of pixels covering each plot area were extracted and an average index value was calculated for each plot for use as predictor variables in multiple linear regression analysis.

Table 4.1. A subset of the spectral indices used for chlorophyll prediction. The name of the index, its formula and the reference when it was first developed is listed for each index.

Index	Formula	Reference
Greenness index	R_{554}/R_{677}	Smith et al., 1995
Modified chlorophyll absorption in reflectance index (MCARI)	$[(R_{700}–R_{670})–0.2*(R_{700}–R_{550})*(R_{700}/R_{670})]$	Daughtry et al., 2000
Modified normalized difference vegetation index (MNDVI)	$(R_{750}–R_{705})/(R_{750} +R_{705})$	Gitelson and Merzlyak, 1996
Lichtenthaler 1st index	$(R_{800}–R_{680})/(R_{800}+R_{680})$	Lichtenthaler et al., 1996
Vogelmann 1st index	R_{740}/R_{720}	Vogelmann et al., 1993
Derivative chlorophyll index (DCI)	$D_{(705)}/D_{(722)}$	Zarco-Tejada et al., 2002
maximum derivative of the red-edge divided by derivative at 703 nm	$D_{max(680-750)})/D_{703}$	Zarco-Tejada et al., 1999
Red-edge inflection point (REIP)	λ_p where $D^2R/D^2\lambda = 0$	Dawson and Curran, 1998

R = reflectance.
D = magnitude of the first derivative.
D^2 = magnitude of the second derivative.
λ = wavelength.

4.2.5. Statistical analyses, model development and testing robustness of models and indices

We examined the temporal variation in the average foliar chlorophyll concentration of all species for a given site and in the foliar chlorophyll concentration for each species. In addition, we wanted to observe the variation in foliar chlorophyll concentration among species within a given a year. This temporal variation could be analyzed using Analysis of Variance (ANOVA), which assumes a normal distribution of the error terms, i.e., the residuals. The Shapiro-Wilk test was used to check for the normal distribution of the residuals (Shapiro and Wilk, 1965). A non-parametric version of ANOVA, the Kruskal-Wallis test, was performed when the residuals were not normally distributed such as the significance of time on average foliar chlorophyll concentration. When time was significant, differences in average foliar chlorophyll concentration among years were tested using the Steel-Dwass multiple comparison test, which is the non-parametric version of Tukey's mean comparison test. To examine the temporal variation in foliar chlorophyll concentration for each species, data were grouped by species and year and an ANOVA run to determine the effect of time on the foliar chlorophyll concentration of trembling aspen, white birch and black spruce, which were sampled across all three years. Since the residuals were not normally distributed for black spruce, non-parametric methods described above were used. Because they were sampled for only two years, white spruce and northern white cedar means were compared using t-tests. Finally, the variation in foliar chlorophyll concentration among species within a given a year was analyzed by ANOVA and Tukey's mean comparison tests or their non-parametric versions.

Indices calculated from the Hyperion remote sensing data of 2005 were used as predictor variables to generate models of chlorophyll concentration in 2005 in multiple regression analysis. We used the best-subsets method, which compares all the possible models with pre-

defined number of predictor variables (Hudak et al., 2006). A maximum of two variables were allowed in the models to avoid overfitting. Variables were kept in the model based on the variance inflation factor (VIF). Generally, a VIF value of 10 or greater has been proposed as an indicator of serious multicollinearity (Kutner et al., 2004). To be more conservative in our analyses, our threshold was a maximum VIF value of 5. Model explanatory power was expressed by the coefficient of determination, R^2 and the adjusted coefficient of determination, R^2_{adj}, which accounts for the improvement in explanatory power of the model as new variables are added. Residuals were tested for normality (Shapiro and Wilk, 1965). Model accuracy was reported as percent of the mean, % RMSE. Models were validated using the leave-one-out cross-validation method. In this method, residuals are calculated by successively leaving one plot out of the analysis and results via the Predicted Residual Sum of Squares (PRESS) RMSE and PRESS r^2. Low values of PRESS RMSE and high values of PRESS r^2 suggest more robust models. The significance of the variables in the model and the model itself was determined based on the p value ($p \leq 0.05$).

Multiple regression with the same criteria as defined above was also run on the 2002 and 2008 data to find common indices for predicting chlorophyll that explained 50% or more of the variation in chlorophyll concentration. The purpose was to identify those indices that are robust in predicting chlorophyll concentration across time and space. In addition, the five two-variable predictive models generated from 2005 data that most accurately predicted chlorophyll concentration were included in the analysis.

Robustness of the models containing one index that could explain 50% or more of the variation in chlorophyll concentration across 2002 and 2005, and the best five two-variable models generated from 2005 data were tested in two ways: (1) the models generated from 2005

were applied to the remote sensing data from 2002, 2004 and 2008 to predict the chlorophyll concentration of that year, and (2) the single and two-variable models were recalibrated based on each year's data and then assessed for how well they predicted chlorophyll concentration for each year.

4.3. Results

Chlorophyll concentration varied significantly among species within a given year as well as among years (Table 4.2). In 2002, differences in chlorophyll concentration between jack pine and the deciduous species were not significant but black spruce had lower chlorophyll concentration than the other three species. In 2005, deciduous species had significantly greater chlorophyll concentration than coniferous species. Among the conifer species, black spruce had the lowest concentration of chlorophyll followed by northern white cedar. In 2008, overlap in chlorophyll concentration among species increased. For example, the chlorophyll concentration of black spruce did not differ significantly from that of the other species. Northern white cedar chlorophyll concentration was significantly lower than that of white spruce and trembling aspen. Unlike the other two years, in 2008 cedar had the lowest chlorophyll concentration and white birch had less chlorophyll than both white and black spruce (the difference between white birch and white spruce was significant). These results indicate that chlorophyll concentration is highly variable among species and from year to year.

Table 4.2. Average foliar chlorophyll concentrations (μg/cm^2) of individual species and all trees. Values are reported as mean ± standard deviation followed by the number of trees sampled for the analysis for a given species in parentheses. Sample size equals 5 for all other species. Uppercase letters denote significant differences in average foliar chlorophyll concentration for a given site (the first row) and for each species among years and lowercase letters denote significant differences in chlorophyll concentration among species within each year at $p \leq 0.05$, respectively.

	2002 (Sudbury)	**2005 (GRFS)**	**2008 (GRFS)**
All trees	36.8 ± 12.3A (n=85)	30.5 ± 10.9B (n=40)	36.8 ± 9.3A (n=50)
Populus tremuloides	44 ± 9.1aA (n=15)	45.9 ± 3.9aA	42.1 ± 7.3abcA (n=10)
Betula papyrifera	40.9 ± 8.9aAB (n=10)	45.4 ± 5.3aA	33.4 ± 5.3adB (n=10)
Picea glauca		30.2 ± 2.6bcA (n=7)	41.9 ± 6.8bB (n=10)
Picea mariana	25.6 ± 7.8bA (n=20)	19.4 ± 4.5dB (n=13)	35.4 ± 11.5abcdC (n=15)
Thuja occidentalis		25.8 ± 1.6cA	27.2 ± 4.3dA
Abies balsamea		34.4 ± 3.4b	
Pinus banksiana	38.6 ± 12.6a (n=40)		

The temporal variation in chlorophyll concentration within species was high for both deciduous and coniferous species. For example, the chlorophyll concentration of trembling aspen and cedar did not vary significantly among years whereas the chlorophyll concentration of black spruce did. Within the same site white birch chlorophyll concentration was significantly lower in 2008 than 2005. On the other hand, white spruce chlorophyll concentration was significantly higher in 2008 than 2005, also within the same site. The average chlorophyll concentration of all trees sampled was identical in 2002 and 2008 but significantly lower in 2005 (Table 4.2).

The indices that could predict chlorophyll with an r^2 value of 0.5 or greater for each year with the exception of 2008 are shown in Table 4.3. The explanatory power of the models was very low for 2008 data with the highest r^2 being 0.11 for single variable models. There were two indices that emerged as good predictors of chlorophyll concentration for temporally and spatially different chlorophyll datasets. These were the DCI and the $D_{max(680-750)})/D_{703}$ indices. These two indices explained 71, 64, 66 and 61, 61, and 62 % of the variation in chlorophyll concentration in 2002, 2004 and 2005, respectively (Table 4.3). When these predictive models derived using 2005 data were used to predict the chlorophyll concentration in all the other years, both models still explained more than half of the variation in chlorophyll concentration except for 2008. The model with the DCI performed better than the $D_{max(680-750)})/D_{703}$ index, which explained 71 and 63 % of the variation compared to 61 and 54 % for 2002 and 2004 chlorophyll data, respectively. Prediction errors ranged from 11.7% for 2002 with the DCI model to 14.6% with the $D_{max(680-750)})/D_{703}$ model for 2004 (Figure 4.2). The plots show separation when grouped by their dominant species. In 2002, trembling aspen and jack pine dominated plots represented the highest values and black spruce dominated plots represented the lowest values with jack pine and a couple of trembling aspen dominated plots constituting the intermediate values of chlorophyll concentration. Trembling aspen dominated and deciduous mixed plots represented the highest values and black spruce dominated and mixed coniferous plots constituted the lowest values of chlorophyll concentration in 2004. For 2008, the same degree of separation between species observed in 2002 and 2004 is not observed. There is more separation with the $D_{max(680-750)})/D_{703}$ index model compared to the DCI index model for the model predicted chlorophyll concentration values but overall the separation is less compared to 2002 and 2004 (Figure 4.2).

Table 4.3. The performance of DCI and $D_{max(680-750)})/D_{703}$ for predicting chlorophyll concentration. PRESS RMSE and PRESS r^2 represent leave-one-out cross-validation error and coefficient of determination, respectively. * implies that the residuals fail to meet the normal distribution assumption.

	index	**r^2**	**Adjusted r^2**	**PRESS RMSE (%)**	**PRESS r^2**	**Residuals Shapiro W and p**
2002	DCI	0.71	0.68	12.6	0.59	0.93, 0.38
	$D_{max(680-750)})/D_{703}$	0.61	0.56	15.1	0.42	0.88, 0.10
2004	DCI	0.64	0.63	13.4	0.59	0.97, 0.42
	$D_{max(680-750)})/D_{703}$	0.61	0.56	6.7	0.49	0.97, 0.45
2005	DCI	0.66	0.65	6.5	0.63	0.96, 0.31
	$D_{max(680-750)})/D_{703}$	0.62	0.61	16.2	0.58	0.98, 0.88
2008	DCI	0.03	<-0.01	7.9	-0.09	0.88, <0.05*
	$D_{max(680-750)})/D_{703}$	0.11	0.08	12.2	0.01	0.86, <0.05*

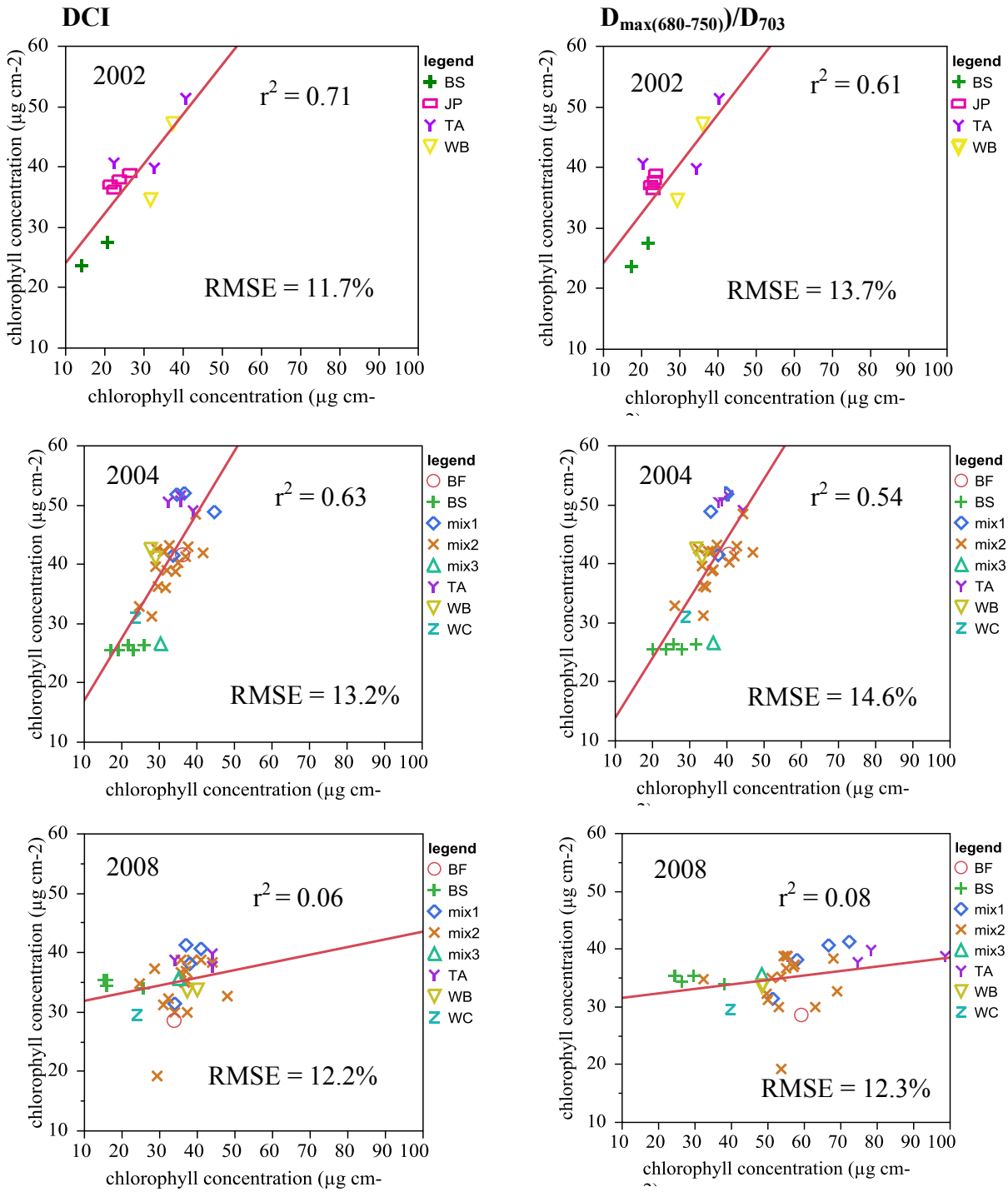


Figure 4.2. Observed versus model predicted chlorophyll concentration using DCI and $D_{max(680-750)})/D_{703}$ indices with the same model parameters as generated from 2005 data (-8.2+28.1*(DCI) & -27.3+39* ($D_{max(680-750)})/D_{703}$). The top graphs show 2005 model predictions for 2002, the middle ones for 2004 and the bottom ones for 2008 chlorophyll concentration, respectively. The prediction RMSEs are reported as percent of the mean of the chlorophyll concentration. BF=balsam fir, BS=black spruce, JP=jack pine, mix1=mixed deciduous, mix2=mixed deciduous and conifer, mix3=mixed conifer, TA=trembling aspen, WB=white birch, WC=white cedar.

The five two-variable models generated from 2005 data that most accurately predicted chlorophyll concentration are shown in Table 4.4. Some of the indices selected were MCARI, Lic 1, Vog 1, $MaxR_{(700-770)}$. The top model had very high VIF values but was retained to assess its performance accuracy for predicting chlorophyll concentration of other sample years. The other four models explained 70% of the variation in chlorophyll concentration, with PRESS RMSEs ranging from 14.3 to 14.8%. When the top model was recalibrated based on 2002 data, the VIF remained high. Neither of the variables was significant in the second model with TCARI/OSAVI and $MaxR_{(700-770)}$ indices. The third model had an R^2 of 0.55 but the PRESS r^2 was 0.02, suggesting it was a weak model. The fourth and fifth models were also not robust because they had 23.7 and 31.2% PRESS RMSE and negative PRESS r^2 values as a result of greater PRESS values than total sum of squares of the model. The first model performed well explaining 71% of the variance in chlorophyll concentration with 13.4% PRESS RMSE and did not have high VIF when 2004 data were used. The other four models each had one insignificant variable. None of the two-variable models predicted 2008 chlorophyll concentration adequately when recalibrated. Neither variable was significant in the model for the first three best models. The fourth and fifth models explained only 5 and 15% of the variation in chlorophyll concentration in 2008.

Table 4.4. The statistics associated with the five best two-variable models generated from 2005 data. PRESS RMSE and PRESS r^2 represent leave-one-out cross-validation error and coefficient of determination, respectively.

Indices in the model	**R^2**	**Adjusted R^2**	**PRESS RMSE (%)**	**PRESS r^2**	**Residuals Shapiro W and p**
MNDVI, Lic 1	0.75	0.73	13.3	0.72	0.96, 0.23
TCARI/OSAVI, $MaxR_{(700-770)}$	0.73	0.71	14.6	0.66	0.98, 0.68
MCARI, $MaxR_{(700-770)}$	0.72	0.70	14.8	0.65	0.98, 0.79
Lic 1, Vog 1	0.72	0.70	14.3	0.67	0.99, 0.98
ZTM, Lic 1	0.72	0.70	14.5	0.66	0.98, 0.85

When the models generated from 2005 data were tested to predict the other years' chlorophyll data with the same parameters, they were not as robust as the single variable models for predicting chlorophyll concentration (Table 4.5). As with the single variable models, the two-variable models also performed very poorly for predicting 2008 chlorophyll data. The greatest amount of variance explained in chlorophyll for 2008 was 15% with the ZTM and Lic 1 model. Models were most accurate for 2004 data. The first model predicted 2004 chlorophyll concentration with 10.8% RMSE and explained 75% of the variation. The second and third models performed very poorly with r^2 values of 0.03. The fourth and fifth models explained 51 and 42% of the variation in chlorophyll concentration with RMSEs of 15.1 and 16.5%, respectively. None of the models could explain more than half of the variation in chlorophyll concentration for 2002. The best model, which included MCARI and $MaxR_{(700-770)}$, explained 35% of the variation in chlorophyll concentration with an RMSE of 17.6% (Table 4.5).

Table 4.5. The predictions of chlorophyll concentration of 2002, 2004 and 2008 by the best two-variable models with the same model parameters as generated from 2005 data. For each model and year, r^2, adjusted r^2 and the RMSE (reported as percent of the mean of the chlorophyll concentration) are listed in the same order, respectively.

	2002	**2004**	**2008**
MNDVI, Lic 1	0.10, <-0.01, 20.7	0.75, 0.74, 10.8	0.06, 0.03, 12.3
TCARI/OSAVI, $MaxR_{(700-770)}$	0.33, 0.25, 17.9	0.03, <-0.01, 21.2	0.02, -0.01, 12.5
MCARI, $MaxR_{(700-770)}$	0.35, 0.27, 17.6	0.03, <-0.01, 21.3	0.03, <-0.01, 12.5
Lic 1, Vog 1	0.23, 0.14, 19.1	0.51, 0.49, 15.1	0.12, 0.09, 11.9
ZTM, Lic 1	0.19, 0.10, 19.6	0.42, 0.38, 16.5	0.15, 0.12, 11.7

4.4. Discussion

4.4.1. Variation in chlorophyll concentration

Chlorophyll concentration varied among years for the spruce and paper birch and among species for a given year. This variation may be due to environmental factors such as stress or

internal regulation. The variation in chlorophyll concentration among species for a given year directly affects the feasibility of generating accurate predictive models for chlorophyll. When variation among chlorophyll concentration of species within a given year is significant, it is possible to obtain more accurate predictive models. Low variation in chlorophyll variation among species makes it difficult to generate accurate predictive models. This is best shown with the 2008 data for which there is a lot of overlap between the chlorophyll concentrations of species. As a result, none of the predictive models explained more than 24% of the variation in chlorophyll concentration despite the increased spatial resolution and higher SNR of AVIRIS data compared to Hyperion. In addition, models generated using 2004 data could not be used to accurately predict chlorophyll concentration in 2008. The seemingly random variation in chlorophyll concentration among species and years makes the goal of continuously assessing boreal mixedwood stand condition and nutritional status via remote sensing difficult using our methods. The gradient created by the presence of multiple species makes it possible to predict chlorophyll concentration at both GRFS and Sudbury sites both with models generated from each years' data and also by using the model generated from 2005 data.

4.4.2. Predictive models

The existence of indices that can predict chlorophyll across time and space is very encouraging. Although originally designed for use at the leaf level, successful application of the DCI and $_{Dmax(680-750))}/D_{703}$ indices to predict chlorophyll concentration at canopy scale is evidence of their robustness. These indices have also been used successfully to estimate leaf and canopy level chlorophyll of sugar maple (*Acer saccharum* Marsh.) using radiative transfer models in Ontario, Canada (Zarco-Tejada et al., 2001; 2002). These two indices were useful for predicting chlorophyll concentration across different growing periods within the same site and on a

different site composed of pure stands of trembling aspen, white birch, black spruce and jack pine. Accuracy levels were similar to the GRFS predictions. Remote sensing data were collected on airborne and spaceborne platforms in these years. More importantly, the models with these two indices calibrated using 2005 satellite remote sensing data perform as well as the models calibrated from the chlorophyll and remote sensing data of each year for GRFS and for different site, the Sudbury site, for 2002. These findings indicate that a universal predictive model for chlorophyll that could be applied to the mixedwood boreal forest is feasible. If this were developed, chlorophyll could be predicted for sites where only remote sensing data are available, i.e., no field measurements are available. However, once developed the model would need to be tested in spatially and structurally varied boreal mixedwood forests. Assuming the test results are promising, future studies could be designed to encompass various boreal types across North America and Eurasia.

The amount of variation in chlorophyll concentration explained and the prediction accuracy of the models were not much improved when the best five two-variable models were examined. Interestingly, neither of the indices selected for single variable models across years were selected for the two-variable models. With the exception of the top model predicting 2004 chlorophyll concentration, the two-variable models were not as robust as those with single indices for predicting chlorophyll when used with the parameters generated from the 2005 data. The top model predicted the 2004 chlorophyll concentration with better accuracy than it did that of 2005. However, this model also had high VIF for 2005 data originally and was retained only to check how it would perform for predicting the chlorophyll concentration of other years.

4.5. Conclusions

In this study, we tested indices for predicting chlorophyll concentration across years for a single site and between sites as well as the robustness of models that could be used to predict chlorophyll concentration. Two indices calculated from the derivative of reflectance in the red-edge (DCI = D_{705}/D_{722} and $D_{max(680-750)})/D_{703}$) emerged as robust predictors of chlorophyll across years and space. Single- variable models were more robust than two-variable models for predicting chlorophyll concentration. The presence of multiple species creates a gradient in the chlorophyll concentration and this gradient makes it possible to predict chlorophyll concentration as well as affecting the performance of predictive models generated using data from a different year. Our results suggest that it may be possible to develop a universal predictive model for chlorophyll concentration that could be applied across the mixedwood boreal forest type but more work is needed to ensure that similar results can be obtained in spatially and structurally different mixedwood types.

4.6. Acknowledgements

The authors gratefully acknowledge the help of Lesley Rich, Denzil Irving, Maara Packalen, Bob Oliver, David Atkinson, Björn Prenzel, Chris Hopkinson, Laura Chasmer, Brock McLeod, Lauren MacLean, and Adam Thompson in the field. Thanks are also due to Al Cameron and Lincoln Rowlinson for establishing a cruise-line trail and for on-site assistance with tree species identification and Garry Koteles for sharing information about tree and understory species and conducting field measurements for the National Forest Inventory (NFI) validation plots. Technical expertise for leaf sampling, processing and laboratory facilities for macronutrient analysis were provided by the Ontario Forest Research Institute (OFRI). Financial assistance for this work was provided by the Department of Forest Resources and Environmental

Conservation at Virginia Tech, a USDA McIntire-Stennis Formula Grant, Queen's University, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canadian Carbon Program (formerly the Fluxnet-Canada Research Network) through funding from NSERC, BIOCAP Canada, and the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS).

4.7. References

Alemdag, I.S., 1983. Mass equations and mechantability factors for Ontario softwoods. Can. For. Serv. Petawawa Natl. For. Inst. Inf. Rep. PI-X-23.

Alemdag, I.S., 1984. Total tree and merchantable stem biomass equations for Ontario hardwoods. Can. For. Serv. Petawawa Natl. For. Inst. Inf. Rep. PI-X-46.

Asner GP, Hicke JA and Lobell DB (2003) Per-pixel analysis of forest structure: vegetation indices, spectral mixture analysis and canopy reflectance modeling. In: Wulder MA and Franklin SE (eds) Remote sensing of forest environments: concepts and case studies. Kluwer, Boston. pp 209–254.

Asner, G. and Martin, R.E. 2008. Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. Remote Sensing of Environment, 112: 3958-3970.

Blackburn, G.A. 1999. Relationships between spectral reflectance and pigment concentrations in stacks of broadleaves. Remote Sensing of Environment 70, 224–237.

Blackburn, G.A. 2007. Hyperspectral remote sensing of plant pigments. Journal of Experimental Botany 58-4: 855-867.

Boardman, J. W., and Kruse, F. A., 1994, Automated spectral analysis: a geological example using AVIRIS data, north Grapevine Mountains, Nevada: in Proceedings, ERIM Tenth Thematic Conference on Geologic Remote Sensing, Environmental Research Institute of Michigan, Ann Arbor, MI, pp. I-407 - I-418.

Chappelle, E. W., Kim, M. S., & McMurtrey, J. E. 1992. Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B and the carotenoids in soybean leaves. Remote Sensing of Environment, 39, 239−247.

Coops, N.C., Stone, C., Culvenor, D.S., Chisholm, L.A., Merton, R.N. 2003. Chlorophyll content in eucalypt vegetation at the leaf and canopy scales as derived from high spectral resolution data. Tree Physiology 23, 23–31.

Curran, P. J., 1989. Remote sensing of foliar chemistry. Rem. Sens. Environ. 30, 271−278.

Curran, J.P., Dungan, J.L., Gholz, H.L., 1990. Exploring the relationship between reflectance red edge and chlorophyll content in slash pine. Tree Physiol. 7, 33-48.

Curran, J.P. 1994. Imaging spectrometry. Progress in Physical Geography 18: 247–266.

Curran, P.J., Kupiec, J.A., Smith, G.M., 1997. Remote sensing the biochemical composition of a slash pine canopy. IEEE Trans. Geosci. Rem. Sens. 35:2, 415-420.

Datt, B. 1998. Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+b, and total carotenoid content in Eucalyptus leaves. Remote Sens. Environ. 66:111–121.

Datt, B., McVicar, T.R., Van Niel, T.G. & Jupp, D.L.B. 2003. Preprocessing EO-1 Hyperion hyperspectral data to support the application of agricultural indexes. IEEE Transactions on Geoscience and Remote Sensing 41-6: 1246-1259.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., Brown De Colstoun, E., McMurtrey, J.E., III. 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment 74: 22–239.

Dawson, T. P. and Curran, P. J. 1998. A new technique for interpolating the reflectance red edge position. International Journal of Remote Sensing, 19, 2133– 2139.

Demming-Adams B, Adams WW. 1996. The role of xanthophyll cycle carotenoids in the protection of photosynthesis. Trends in Plant Science 1, 21–27.

Elvidge, C.D., Chen, Z. 1995. Comparison of broad-band and narrow-band red and near-infrared vegetation indices. Remote Sensing of Environment 54: 38-48.

ENVI. 2009. Atmospheric Correction Module: QUAC and FLAASH User's Guide. Version 4.7.

Filella, I., Peñuelas, J., 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. Int. J. Rem. Sens. 15:7, 1459-1470.

Gitelson, A. A., and Merzlyak, M. N. 1996. Signature analysis of leaf reflectance spectra: Algorithm development for remote sensing of chlorophyll. *J. Plant Physiol.* 148:494–500.

Gitelson, A.A., Viña, A., Verma, S.B., Rundquist, D.C., Arkebauer, T.J., Keydan, G., Leavitt, B., Ciganda, V., Burba, G.G., Suyker, A.E. 2006. Relationship between gross primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity, Journal of Geophysical Research 111, D08S11, doi:10.1029/2005JD006017.

Gobron, N., Pinty, B. and Verstraete, M.M. 1997. Theoretical limits to the estimation of the leaf area index on the basis of visible and near-infrared remote sensing data. IEEE Transactions on Geoscience and Remote Sensing 35: 1438–1445.

Green, A. A., Berman, M., Switzer, P., and Craig, M. D. 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and Remote Sensing 26: 65– 74.

Hendry, G.A.F., J. D. Houghton, and S.B. Brown. 1987. Tansley review no. 11: the degradation of chlorophyll—a biological enigma. New Phytologist 107: 255–302.

Horler, D.N.H., Dockray, M., Barber, J., Barringer, A.R., 1983. Red edge measurements for remotely sensing plant chlorophyll content. Adv. Space Res. 3, 273–277.

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P.E. & Morgan, P. 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return LiDAR and multispectral satellite data. *Canadian Journal of Remote Sensing* 32: 126–138.

Kutner, M., Nachtsheim, C., and Neter, J. 2004. Applied Linear Regression Models. 4th Ed. McGraw-Hill.

le Maire G., Francois C., Dufrene E. 2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. Remote Sensing of Environment 89, 1–28.

Lichtenthaler, H.K., Gitelson, A.A. and Lang, M.1996 Non-destructive determination of chlorophyll content of leaves of a green and an aurea mutant of tobacco by reflectance measurements. Journal of Plant Physiology, 148, 483–493.

Peñuelas, J., Baret, F., & Filella, I. 1995. Semiempirical indices to assess carotenoids chlorophyll: A ratio from leaf spectral reflectance. Photosynthetica, 31, 221−230.

Shapiro, S.S. & Wilk, M.B. 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52-3/4: 591-611.

Sims, D. A., & Gamon, J. A. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sensing of Environment, 81, 337−354.

Smith, R. C. G., Adams, J., Stephens, D. J., and Hick, P. T. 1995. Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. Australian Journal of Agricultural Research, 46, 113– 125.

Stagakis, S., Markos, N., Sykioti, O., Kyparissis, A. 2010. Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a *Phlomis fruticosa* Mediterranean ecosystem using multiangular CHRIS/PROBA observations. Remote Sensing of Environment, 114, 977-994.

Thomas, V., Treitz, P., McCaughey, J.H., Noland, T., Rich, L., 2008. Canopy chlorophyll concentration estimation using hyperspectral and lidar data for a boreal mixedwood forest in northern Ontario, Canada. Int. J. Remote Sens. 29, 1029–1052.

Ustin, S. L., Gitelson, A. A., Jacquemoud, S., Schaepman, M. E., Asner, G., Gamon, J. A., & Zarco-Tejada, P. 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sensing of Environment, 113 (suppl. 1), S67−S77.

Vogelmann, J.E., Rock, B.N. & Moss, D.M. 1993. Red edge spectral measurements from sugar maple leaves. *International Journal of Remote Sensing* 14: 1563–1575.

Wellburn, A.R., 1994. The spectral determination of chlorophylls a and b as well as total carotenoids using various solvents with spectrophotometers of different resolutions. J. Plant Physiol. 144, 307–313.

Wu, C., Niu, Z., Tang, Q., Huang, W., Rivard, B., Feng, J. 2009. Remote estimation of gross primary production in wheat using chlorophyll-related vegetation indices. Agricultural and Forest Meteorology, 149: 1015-1021.

Wu, C., Xiuzhen, H., niu, Z. Dong, J. 2010. An evaluation of EO-1 hyperspectral Hyperion data for chlorophyll content and leaf area index estimation. International Journal of Remote Sensing 31-4: 1079-1086.

Zarco-Tejada, P.J., Miller, J.R., Mohammed, G.H., Noland, T.L., Sampson, P.H. 1999. Canopy optical indices from infinite reflectance and canopy reflectance models for forest condition monitoring: application to hyperspectral CASI data. In IEEE 1999 International Geoscience and Remote Sensing Symposium, IGARSS ’99, 28 June–2 July, (Hamburg: IEEE); 3: 1878–1881.

Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H., Sampson, P.H., 2001. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll estimation in closed forest canopies with hyperspectral data. IEEE Trans. Geosci. Rem. Sens. 39, 1491–1507.

Zarco-Tejada, P.J., Miller, J.R., Mohammed, G.H., Noland, T.L., Sampson, P.H. 2002. Vegetation stress detection through chlorophyll a+b estimation and fluorescence effects on hyperspectral imagery. Journal of Environmental Quality 31: 1433-1441.

Zarco-Tejada, P.J., Miller, J.R., Harron, J., Baoxin. H., Noland, T.L., Goel, N., Mohammed, G.H., Sampson, P.H. 2004. Needle chlorophyll content estimation through model inversion using hyperspectral data from boreal conifer forest canopies. Remote Sensing of Environment 89, 189-199.

Zhang, Q., Middleton, E.M., Margolis, H.A., Drolet, G.G., Barr, A.A., Black, T.A. 2009. Can a satellite-derived estimate of the fraction of PAR absorbed by chlorophyll (FAPARchl) improve predictions of light-use efficiency and ecosystem photosynthesis for a boreal aspen forest? Remote Sensing of Environment, 113: 880-888.

Chapter 5 Conclusions

5.1. Summary and findings

The results of this study have shown that spaceborne IS data (Hyperion) can be used to predict 1) the N:P ratio, which is an indicator of nutrient limitation, 2) macronutrients including N, P, K, Mg, and 3) chlorophyll in a boreal mixedwood forest canopy. The ability to model the N:P ratio and macronutrients using spaceborne Hyperion data demonstrates the potential for mapping them at the canopy scale across larger geographic areas with lower cost. These maps would be useful for understanding forest physiology, nutritional status and health. They can also be used to estimate ecosystem processes such as GPP or as input layers to increase the accuracy of ecosystem process models such as PnET. In addition, LiDAR data provided the only accurate estimate of another macronutrient, Ca and more improved predictions of K and Mg in this boreal forest environment. The relationship of LiDAR data to macronutrient concentration is through relationships to canopy structure, specifically canopy height and crown closure which are in turn correlated with the macronutrients.

More detailed findings specific to each objective are presented below:

Objective 1: to estimate the canopy N:P ratio using airborne and spaceborne IS data, address the effects of temporal and spatial resolution variation on prediction accuracy of the canopy N:P ratio and investigate whether LiDAR data have any explanatory power for the estimation of the canopy N:P ratio.

Narrowband indices calculated from airborne and spaceborne IS data explained very similar amount of variance in the canopy N:P ratio. The index originally selected from airborne IS data was temporally and spatially more robust. Although LiDAR models did not provide

improved explanatory power or prediction accuracy for predicting the canopy N:P ratio over IS data, they provide insights into the relationships of structure with growth and productivity at the site. N:P ratio is a measure of nutrient limitation of biomass production and therefore is directly related to growth and productivity. The variation in growth as manifested by variation in height is being captured by the LiDAR metrics. The results suggest that the canopy N:P ratio can be predicted by remote sensing data based on the relationship between the canopy N:P ratio and crown closure at this site. The presence of multiple species creates a gradient in crown closure and thus makes it feasible to predict the N:P ratio using remote sensing. Variation in spatial resolution resulted in significant differences in the N:P ratio prediction and temporal variation affects the prediction results based on the degree of gradient and the variation present within a given season's N:P ratio data. These results are encouraging as they show that the spatially-explicit modeling of the canopy N:P ratio is possible by remote sensing and this can provide a diagnostic tool for assessing nutrient limitation.

Objective 2: to evaluate the utility of spaceborne IS data to estimate canopy macronutrient (N, P, K, Ca and Mg) concentrations in a mixedwood boreal forest; and test the potential contribution of canopy structural information derived from LiDAR data for improving the predictive models for these macronutrients.

Spaceborne Hyperion data could explain the variance ranging from 64 to 75% of the variance in N, P, K, Mg concentrations. LiDAR data improved the prediction of K and Mg over Hyperion data alone and Ca could only be accurately predicted by LiDAR data. Ca is an element that requires canopy structure data for accurate prediction at this study site. On the other hand, P could be predicted by Hyperion data the most accurately. Finally, the predictive ability for K was

improved when the IS and LiDAR data were combined. The presence of multiple species at the site is driving the relationships between the LiDAR data and macronutrient concentrations by creating a gradient.

Objective 3: to identify spectral indices that can predict canopy chlorophyll concentration across years, test the robustness of the models with these indices derived from spaceborne IS data for predicting canopy chlorophyll concentration across years within the same site, and extend these predictive models for canopy chlorophyll concentration to a different geographic location and time.

Two indices calculated from the derivative of reflectance in the red-edge (DCI = D_{705}/D_{722} and $D_{max(680-750)})/D_{703}$) emerged as robust predictors of chlorophyll across years and space. These two indices explained 63 and 54, and 71 and 61% of the variance in chlorophyll concentration when applied for prediction of chlorophyll across time, space and sensors. Two-variable models were not as robust for predicting chlorophyll as single variable models. The presence of multiple species creates a gradient in the chlorophyll concentration and this gradient makes it possible to predict chlorophyll concentration as well as affecting the performance of predictive models generated using data from a different year. The results suggest that there is a possibility of a universal predictive model for chlorophyll that could be applied across the mixedwood boreal forest type if similar results can be obtained in spatially and structurally different mixedwood boreal forest types.

These findings suggest that LiDAR data do not contribute significantly to the prediction of macronutrients or the N:P ratio with the exception of Ca at this boreal mixedwood forest. As a result, it would not be the ideal remote sensing data type for macronutrient prediction since it is

still rather expensive to acquire airborne LiDAR data. Spaceborne IS data performed well for predicting the N:P ratio and macronutrients. The explanatory power of the predictive models obtained from airborne and spaceborne IS data were very similar. Models generated from spaceborne data at one site also perform well at a different mixedwood boreal site as long as the gradient in the foliar biochemical exists. The interspecific variation in the foliar biochemistry concentration within a given year and how that varies from year to year has a direct impact on the ability to generate predictive models, regardless of the IS data type, i.e., airborne vs. spaceborne. This is related to the statistical approach used in the study, i.e., regression, by which models are generated based on the variation and the gradient in the response variable. Based on the degree of interspecific variability over time, it was not possible to generate predictive models every year within the same site. In this context, process-based approaches which describe the interaction of leaves and canopies with light may be better suited to overcome these problems despite their complexity compared to the empirical methods. The canopy concentration maps of macronutrients provide spatially continuous and high spatial resolution information about the nutritional and physiological status of this forest, which can not be obtained from a species map. Additionally, the spatial distribution of macronutrient concentration tends to parallel with the species distribution at this site where those areas dominated by the deciduous species have higher concentrations of macronutrients and the areas dominated by coniferous species have lower concentrations. This brings about the possibility of species mapping based on foliar biochemistry data if separation among tree species exists when two or more macronutrients are used (Figure 5.1). However, this separation of species may not be possible in other forest ecosystems such as the temperate forest, which has high species diversity unlike the boreal forest.

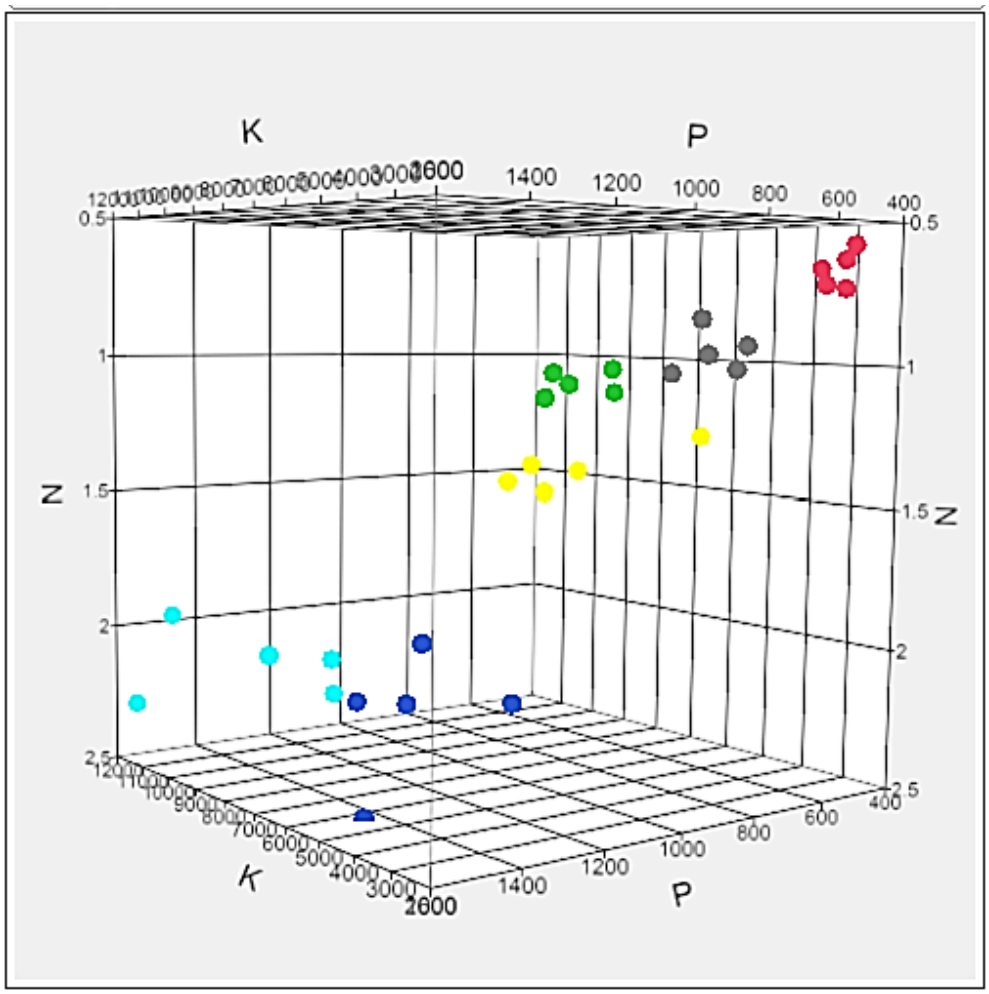


Figure 5.1. Scatterplot of foliar N, P and K samples of the six species found at the GRFS (color coded) in 3-D space.

5.2. Limitations and future directions

The study is small in scope. It was conducted in the footprint of a flux tower with 1km radius covering an area of approximately 314 hectares (with the exception of Chapter 4, which includes a study area near Sudbury, Ontario). It is representative of one type of boreal forest, the mixedwood forest. Therefore, generalizations to other types of boreal forests cannot be made. The method used in for the analysis, i.e., simple and multiple linear regression, is empirical and suffers from the common deficiencies identified with empirical models such as the need for

recalibration with data when applied at a different time or in a different location. Additionally, correlation does not imply causation, so the analyst needs to make sure in advance that the predictors used in the analysis have a causal relationship with the response variable. The success of generating a predictive model is profoundly dependent on the presence of a gradient in the concentration of the biochemical being estimated. If there is overlap between species, it is not possible to generate predictive models regardless of the quality of the remote sensing data. Temporal variation is another impediment to the generation of predictive models. A model obtained from low SNR satellite data which is robust in predicting chlorophyll for a different site at a different time does not perform well at another time even though the remote sensing data source is airborne with higher SNR because there is too much overlap among species in chlorophyll concentration in that particular year, which was the case for chlorophyll concentration distribution at GRFS in 2008. GRFS is a mixedwood site, with coniferous and deciduous species occurring together and thus creating the gradient necessary to generate models. Prediction at sites with single species might present difficulties due to low variation in the biochemical of interest unless a gradient exists due to some treatment such as fertilization. Even though Hyperion data are available for free, there is no continuous coverage making it difficult to obtain data for large areas. Also, the SNR of Hyperion data is lower compared to airborne data and requires significant preprocessing. Because of the above mentioned reasons, the findings of this study are experimental or at maximum useful for small patches of forests. An operational canopy biochemistry estimation approach will require continuous availability of IS data with high SNR and well-tested and robust methods. The limitations associated with the remote sensing part of the problem might be addressed by the launch of NASA's HyspIRI mission or the German Aerospace Center's EnMAP, both of which will provide global

spaceborne IS data with high spectral and temporal resolution. Both of these missions are scheduled to be launched within this decade.

Future studies should test the feasibility of the estimation of the N:P ratio and macronutrients in other types of the boreal forest as well as in different ecosystems of the Earth. The robustness of models also needs to be investigated by applying them to spatially and temporally different datasets. If model(s) that could be applied for a given ecosystem type is identified, then the prediction of a given biochemical for the whole ecosystem would be simplified greatly where only remote sensing data covering the distributional range of the ecosystem could be used to apply the models. This in turn would pave the way for the production of spatially more accurate global datasets (e.g. a global chlorophyll concentration map) which could be used in global assessments of productivity.