

Canopy nitrogen: a remote sensing and modelling approach

**Stikstof in het bladerdak:
een aardobservatie en modelleer benadering**

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 29 juni 2021 des middags te 2.15 uur

door

Yasmina Mathilde Amélie Loozen

geboren op 19 december 1989
te Luik, België

Promotoren:

Prof. dr. S.M. De Jong

Prof. dr. M.J. Wassen

Prof. dr. D.J. Karssenber

Copromotor:

Dr. K.T. Rebel

This thesis was supported by the Netherlands Organization for Scientific Research (NWO) under the grant number NWO ALW-GO-AO/14-12.

Utrecht Studies in Earth Sciences 232

Canopy nitrogen: a remote sensing and modelling approach

Yasmina Loozen

Utrecht 2021

ISBN 978-90-6266-597-6
ISSN 2211-4335 (USES Series)

Correspondence to: yasmina.loozen@gmail.com

Cartography, figures and layout: C&M (9914), Faculty of Geosciences, Utrecht University
Cover: Carolina Levicek, CΠG

Printed in the Netherlands by Ipskamp

This work is licensed under the Creative Commons
Attribution 4.0 International Licence, <http://creativecommons.org/licenses/by/4.0/>

Contents

Summary	9
Nederlandse samenvatting	10
Acknowledgements	12
Acronym list	14
Chapter 1 Introduction	15
1.1 Relevance	16
1.2 The carbon cycle and the terrestrial biosphere	16
1.3 N cycle and canopy N	18
1.4 Vegetation models	19
1.5 Canopy N remote sensing	20
1.6 Research questions and outline of the thesis	22
1.7 References	24
Chapter 2 Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses	33
2.1 Introduction	35
2.2 Material and methods	37
2.2.1 Culture of the plants	37
2.2.2 Reflectance measurements	37
2.2.3 Leaf chemical measurements	38
2.2.4 Spectral bands considered: original and resampled to satellite bands	38
2.2.5 Data analysis	39
2.3 Results	41
2.3.1 Descriptive statistics of canopy N:P, canopy N and canopy P	41
2.3.2 Original narrow band spectra	41
2.3.3 Spectra resampled to satellite sensors' bands	45
2.4 Discussion	45
2.4.1 Original narrow band spectra	45

2.4.2	Spectra resampled to satellite sensors' bands	46
2.4.3	Future perspectives	47
2.5	Conclusion	47
2.6	Acknowledgments	48
2.7	Appendix	48
2.8	References	58

Chapter 3	Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index	63
3.1	Introduction	65
3.2	Material and methods	67
3.2.1	Study area	67
3.2.2	Data collection	69
3.2.2.1	Canopy N data	69
3.2.2.2	MTCI product	71
3.2.3	Data handling	72
3.2.3.1	Methodology to link canopy N data to MTCI values	72
3.2.3.2	Relationship between MTCI and canopy N data at lower spatial resolution	72
3.2.3.3	Relationship between MTCI and canopy N data at original spatial resolution	73
3.2.3.4	Statistical analysis	73
3.3	Results	74
3.3.1	Descriptive statistics	74
3.3.2	Relationship between MTCI and canopy N data at lower spatial resolution	75
3.3.3	Relationship between MTCI and canopy N data at original spatial resolution	77
3.3.3.1	Relationship between MTCI and canopy N concentration	77
3.3.3.2	Relationship between MTCI and canopy N content	78
3.4	Discussion	79
3.4.1	Relationship between MTCI and canopy N data at lower spatial resolution	79
3.4.2	Relationship between MTCI and canopy N data at original spatial resolution	79
3.4.2.1	Canopy N concentration	79
3.4.2.2	Canopy N content	80
3.4.2.3	Comparing results obtained for canopy N concentration and canopy N content	81
3.4.3	Possible confounding factors of the MTCI canopy N relationship	81
3.4.4	Perspectives for future applications	82
3.5	Conclusion	83
3.6	Data availability	83
3.7	Acknowledgements	84
3.8	Appendix	84
3.9	References	91

Chapter 4 Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method	97
4.1 Introduction	99
4.2 Material and methods	101
4.2.1 Canopy N data	101
4.2.1.1 ICP Forests	101
4.2.1.2 Canopy N data analysis	101
4.2.2 Environmental variables	101
4.2.2.1 Bioclimatic variables	101
4.2.2.2 Altitude	102
4.2.2.3 Soil properties	102
4.2.2.4 N deposition	102
4.2.2.5 Land cover	102
4.2.3 Remote sensing variables	103
4.2.3.1 MOD13Q1 product	103
4.2.3.2 MTCI product	103
4.2.4 Data preprocessing	103
4.2.5 Random forests	104
4.2.6 Mapping canopy N	106
4.3 Results	106
4.3.1 Descriptive analysis of canopy N plot data	106
4.3.2 Results of the random forests analysis	109
4.3.3 Variable importance	111
4.3.4 Canopy N map for European forests	113
4.4 Discussion	115
4.4.1 Canopy N spatial pattern	115
4.4.2 Comparison with published studies	115
4.4.3 The role of environmental variables	115
4.4.4 Variables importance	116
4.4.5 Source of errors	117
4.4.6 Future perspectives	117
4.5 Conclusion	118
4.6 Acknowledgements	118
4.7 References	119
Chapter 5 Comparison of simulated foliage nitrogen by the O-CN and LPJ-GUESS vegetation models with a canopy nitrogen map based on forest sampling	125
5.1 Introduction	127
5.2 Material and methods	128
5.2.1 Global vegetation models	128
5.2.1.1 O-CN	128

5.2.1.2 LPJ-GUESS	129
5.2.2 Random forests map	130
5.2.3 Data preprocessing	130
5.3 Results	131
5.3.1 Spatial patterns in canopy N	131
5.3.2 Relationships between modelled canopy N values	136
5.4 Discussion	136
5.4.1 Canopy N spatial pattern	136
5.4.2 Outlook on future research	138
5.5 Conclusion	139
5.6 References	140
Chapter 6 Synthesis	145
6.1 Context	146
6.2 Remotely sensed vegetation indices for canopy N estimation	146
6.3 Including environmental variables as predictors	149
6.4 Remote sensing canopy N estimates for global vegetation models	151
6.5 Future perspectives	152
6.6 References	154
Peer-reviewed publications	158
Curriculum vitae	159

Summary

In the last decades, CO₂ emissions from fossil fuel burning and land use change have caused an increase in the CO₂ concentration of the atmosphere. This human induced rise in atmospheric CO₂ concentration is extremely likely to be causing climate change, which poses a threat to human communities worldwide. The increase in atmospheric CO₂ influences the terrestrial biosphere. Over the last decade, the terrestrial biosphere acted as a carbon sink, absorbing an estimated 29% of the yearly anthropogenic CO₂ emissions. The behavior of the terrestrial C sink with future CO₂ emission is, however, uncertain. Nitrogen (N) is an essential and limiting nutrient for vegetation growth. Due to the resulting close links between the carbon (C) and N cycles, the N cycle influences the biosphere response to the rise in atmospheric CO₂. How the N cycle influences the C cycle in terrestrial biosphere is studied with global vegetation models. Data on the N cycle are needed for the models. However, data on the N cycle are lacking at global scale. Remote sensing methods could provide insights on the spatial pattern of canopy N, defined as the foliar N scaled to the whole vegetation canopy, at large scale. In this context, this thesis explores the possibility to estimate canopy N concentration across scales and vegetation types. This is done under different conditions using vegetation indices (VIs) computed from remotely sensed spectral reflectance. In the first study (*Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses*), the reflectance spectra of a grass species, *Holcus lanatus* L., was measured under controlled conditions in a laboratory experiment. Several VIs were subsequently computed and related to the grass canopy N concentration using linear regressions. In the second study (*Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index*), canopy N was estimated in a Mediterranean forest ecosystem at regional scale in Catalonia, Spain, using the MERIS Terrestrial Chlorophyll Index (MTCI) obtained from the MERIS sensor aboard ESA-Envisat satellite. In the third study (*Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method*), canopy N was estimated in European forests at continental scale using either only satellite-based variables or both VIs and environmental variables as predictors with the random forests method algorithm. In the last study (*Comparison of simulated foliage nitrogen by the O-CN and LPJ-GUESS vegetation models with a canopy nitrogen map based on forest sampling*), the obtained canopy N map in European forests was compared with foliage N simulated by two global vegetation models, O-CN and LPJ-GUESS, in Europe. To conclude, the results obtained in this thesis showed that remotely sensed VIs could be related to canopy N from different vegetation types, i.e. grasses, Mediterranean forests and European forests, across spatial scales. These results could lead to the use of VIs to map canopy N globally provided sufficient ground-truth data are available.

Nederlandse samenvatting

In de afgelopen decennia hebben CO₂-emissies ten gevolge van de verbranding van fossiele brandstoffen en veranderingen in landgebruik geleid tot een toename van de CO₂-concentratie in de atmosfeer. Deze door de mens veroorzaakte stijging van de atmosferische CO₂-concentratie is zeer waarschijnlijk de oorzaak van de huidige klimaatverandering. De toename van de atmosferische CO₂-concentratie is van invloed op de terrestrische biosfeer. In het afgelopen decennium werd er in de terrestrische biosfeer koolstof opgeslagen. De netto opname van terrestrische ecosystemen bedroeg ongeveer 29% van de jaarlijkse antropogene CO₂-uitstoot. Het is echter onzeker of de terrestrische biosfeer zich gelijk blijft gedragen onder de toekomstige CO₂-uitstoot. Stikstof (N) is een essentiële en limiterende voedingsstof voor de groei van vegetatie. Daardoor zijn de koolstof (C) cyclus en de N cyclus nauw met elkaar verbonden, en beïnvloedt de N-cyclus de reactie van de biosfeer op de stijging van atmosferische CO₂-concentratie. De relaties en wisselwerkingen tussen de N-cyclus en de C-cyclus in de terrestrische biosfeer worden vaak bestudeerd met behulp van mondiale vegetatiemodellen. Voor deze modellen zijn gegevens nodig over de N-cyclus. Echter, gegevens over de N-cyclus op mondiale schaal ontbreken. Methoden uit de aardobservatie kunnen inzicht verschaffen in het grootschalige ruimtelijke patroon van N in het bladerdak, gedefinieerd als het percentage N in het bladerdak van de vegetatie. In deze context wordt in dit proefschrift de mogelijkheid onderzocht om de ruimtelijke variatie in N-concentratie in het bladerdak te bepalen, op verschillende ruimtelijke schaalniveaus en voor verschillende vegetatietypes. Hierbij wordt gebruik gemaakt van vegetatie-indices (VI's) die worden berekend op basis van de spectrale reflectie gemeten door middel van aardobservatie. In de eerste studie (*Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses*) werden de reflectiespectra van een grassoort, *Holcus lanatus* L., gemeten onder gecontroleerde omstandigheden in een laboratoriumexperiment. Vervolgens werden verschillende VI's berekend en gerelateerd aan de N-concentratie in de bladeren van de grassen met behulp van lineaire regressies. In de tweede studie (*Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index*) werd de hoeveelheid stikstof geschat in het bladerdak van een mediterrane boscossysteem op regionale schaal in Catalonië, Spanje, met behulp van de MERIS Terrestrial Chlorophyll Index (MTCI), verkregen met de MERIS-sensor aan boord van de ESA-Envisat-satelliet. In de derde studie (*Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method*) werd de hoeveelheid stikstof in het bladerdak van Europese bossen op continentale schaal bepaald. Hierbij werd een vergelijking gemaakt tussen een methode alleen gebaseerd op gegevens van aardobservaties en een methode waarbij ook aanvullende ruimtelijke omgevingsvariabelen werden gebruikt. Deze gegevens werden gebruikt als ruimtelijke voorspellers in het Random Forests machine learning algoritme, resulterend in een stikstofkaart op Europese schaal. In de laatste studie

(Comparison of simulated foliage nitrogen by the O-CN and LPJ-GUESS vegetation models with a canopy nitrogen map based on forest sampling) werd deze Europese stikstofkaart vergeleken met bladstikstof die door twee mondiale vegetatiemodellen, O-CN en LPJ-GUESS, in Europa werd gesimuleerd. De resultaten van mijn proefschrift laten zien dat vegetatie indices ' die uit aardobservaties verkregen zijn, gerelateerd kunnen worden aan stikstof in het bladerdak van verschillende vegetatietypen, zoals graslanden, mediterrane bossen en andere Europese bossen, op verschillende ruimtelijke schaalniveaus. Deze resultaten dragen bij aan pogingen om met gebruikmaking van vegetatie indices stikstof in het bladerdak wereldwijd in kaart te brengen, op voorwaarde dat er voldoende data van metingen op de grond beschikbaar zijn.

Acknowledgements

I feel really grateful I had the opportunity of undergoing this PhD experience. During this time at Utrecht university, I have grown and learned so much, certainly as a researcher but also as a person. Utrecht is a beautiful city, I feel blessed I could call it my home for four years. A piece of my heart stayed in Utrecht when I left the Netherlands. This experience would not have been as rich and fulfilling without all the people I had the opportunity to meet and get to know.

First, I would like to thank my supervisors for believing in me and giving me the chance of pursuing this PhD. Although I did not know any of them before starting the PhD, I quickly felt really fortunate to be mentored by such involved and available persons. I would not have been able to reach the end of the PhD without their guidance and support. Steven and Martin, thank you for bringing your experience and knowledge in your respective fields to my research project. Your comments and suggestions during our meetings and on the many versions of the article manuscripts were often helping me seeing the scientific challenges from a different perspective. Steven, thank you also for your support towards the end of the PhD to ensure I could finalize the thesis. I met with Karin and Derek every week for four years and most of the directions of the PhD were decided during these meetings. Derek, thank you for your incisive and thoughtful comments, for sharing some of your expertise in spatial statistics and modelling and your continuous support. Karin, thank you for your passion for the research subject, your patience with me and your incredible positive energy you brought to the project.

Then, what comes to my mind when I think about this PhD adventure are the colleague I shared a room with, Floris, Mart and Judith and then John, Myrna and Ineke. Floris, Mart, Ineke, Judith and Myrna, thank you for making me discover the Dutch culture and answering my sometimes awkward questions about how you do things in the Netherlands. To all, I enjoyed the uplifting talks, the suggestions regarding the work and the relaxing moments we shared together, thank you for that!

The PhD would have been different without the colleagues from the Copernicus institute on the 11th floor of the Unnik building and then the 8th floor of the new building, Jetske, Elizabeth, Rémon, Mara, Maria, Jerry, Shuqiong, Svenja, Iris, Leontien, Dona, Ana, Ann-Hélène, Koen, Jiefei, Brian, Annick, Adrien, Vincent, Floris, Feroz, Gillian, David, Max, Stefan, Hugo, Maarten and all the people I'm certainly forgetting, thank you for the nice atmosphere, the chats around the coffee corner, checking in with the advancement of the PhD, sharing your experience as PhD candidates and researchers, the inspiring discussions, the conferences we went together, the diners we shared at the Chinese restaurant, the new year's parties, I enjoyed these moments and they made me grow. They definitely contributed to make my time in Utrecht a nice memory.

I also would like to acknowledge the teams of the physical geography lab and the greenhouse garden for their support during the spectrometer experiment as well as Ton Markus and Margot Stoete, for their help with the figures and design of the thesis. I am also grateful to Maarten Zeilmans, Oliver Schmitz and Meng Lu from the physical geography department for sharing their expertise with me.

I would like to thank my new colleagues of the Remote Sensing and Geodata Unit at ISSeP for their encouragements and their flexibility regarding my PhD work, allowing me to reach the finishing line.

À ma famille, Malaïka, Brigitte, Baudouin, Franck, Amandine, Malcom et Émeline, merci pour les encouragements et le soutien que vous avez montré pendant ces années. Merci de m'avoir permis d'avancer sur le travail en gardant Zola, sans quoi je n'aurais certainement pas pu voir la fin de ce doctorat.

À mes parents, merci de m'avoir toujours soutenue dans mes différentes envies et projets. Sans votre investissement et confiance en amont, je n'aurais certainement pas pu prendre part à cette expérience, ni la terminer.

Romain, merci pour ta patience, ton soutien et ton amour toutes ces fois où j'étais convaincue que je n'allais pas y arriver. Pour ton intérêt sincère pour le sujet de recherche aussi.

Zola, merci d'être la merveilleuse personne que tu es, pour ta lumière et merci de me montrer à quel point je peux être forte et résiliente. You make this world a better place!

Acronym list

C	Carbon
DBH	Diameter at breast height
ESA	European Space Agency
EVI	Enhanced vegetation index
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
VIs	Vegetation indices
ENF	Evergreen broadleaf forest
DBF	Deciduous broadleaf forest
EBF	Evergreen broadleaf forest
FPC	Fractional PFT cover
GVM	Global vegetation model
MSE	Mean squared error
MTCI	MERIS Terrestrial chlorophyll index
N	Nitrogen
NIR	Near-infrared
NDVI	Normalized difference vegetation index
OOB	Out-of-bag
P	Phosphorus
PFT	Plant functional type
RRMSE	Relative root mean squared error
RS	Remote sensing

Chapter 1

Introduction

1.1 Relevance

In the last decades, CO₂ emissions from fossil fuel burning and land use change have caused an increase in the CO₂ concentration of the atmosphere (Friedlingstein et al. 2019). This led global CO₂ concentrations to reach unprecedented high levels compared to the last 800,000 years, as inferred from ice core measurements (Ciais et al. 2013). This human induced rise in atmospheric CO₂ concentration is causing climate change (Bindoff et al. 2013; Cook et al. 2013; Cook et al. 2016). Compared to the reference period of 1850-1900, the mean global temperature of 2006-2015 has increased by about 0.87 °C (0.75 °C and 0.99 °C) (Allen et al. 2018). Other consequences include a higher frequency and intensity of extreme events such as heat-waves, drought events and heavy precipitation (Jia et al. 2019) as well as land degradation (Olsson et al. 2019) and desertification (Burrell et al. 2020; Mirzabaev et al. 2019). This has negative impacts on food security and ecosystems, on which human society relies for numerous ecosystem services (Mbow et al. 2019). Due its threat to human communities, urgent action to combat climate change and its impacts has been identified as one of the sustainable development goals of the United Nations (United Nations 2020).

In this context, it is necessary to understand the responses of the biosphere to the increase in atmospheric CO₂. Nitrogen (N) is essential for vegetation growth. Due to the resulting close links between the carbon (C) and N cycles, the N cycle will influence the biosphere response to the rise in atmospheric CO₂. By improving our understanding of the N cycle, and by quantifying spatial patterns in canopy N, we will also enhance our understanding of the biosphere response to future changes in the C cycle.

This thesis aims at exploring approaches to remotely sense N concentration in plant foliage. This is done at different scales and in different ecosystems. Scientifically, this work examines if it is feasible to apply remote sensing methods to map spatial patterns of foliage N at regional to continental scales. Results are evaluated against observations of canopy N as well as canopy N values estimated by terrestrial vegetation models. By doing so, the research presented here will contribute to the improvement of vegetation models, which are essential in understanding the consequences of the changes in the C cycle in terrestrial systems, both at present time and in the future. Ultimately, this work will also contribute to improve our scientific understanding of the N and C cycles.

1.2 The carbon cycle and the terrestrial biosphere

Atmospheric CO₂ has risen globally since pre-industrial times because of anthropogenic activities, with fossil fuel emissions being the dominant source of CO₂ to the atmosphere (Friedlingstein et al. 2019). In 2018, the atmospheric CO₂ concentration reached 407 ppm (Friedlingstein et al. 2019). This human induced increase in atmospheric CO₂ concentration is extremely likely to cause climate change (Bindoff et al. 2013; Cook et al. 2013; Cook et al. 2016), which resulted in a 0.87 °C (0.75 °C and 0.99 °C) rise in mean global temperature in 2006-2015 compared to the 1850-1900 period (Allen et al. 2018).

Atmospheric CO₂ is coupled with the terrestrial biosphere. Over the last decade, an estimated 29% of the yearly anthropogenic CO₂ emissions was absorbed by the terrestrial

The global carbon cycle

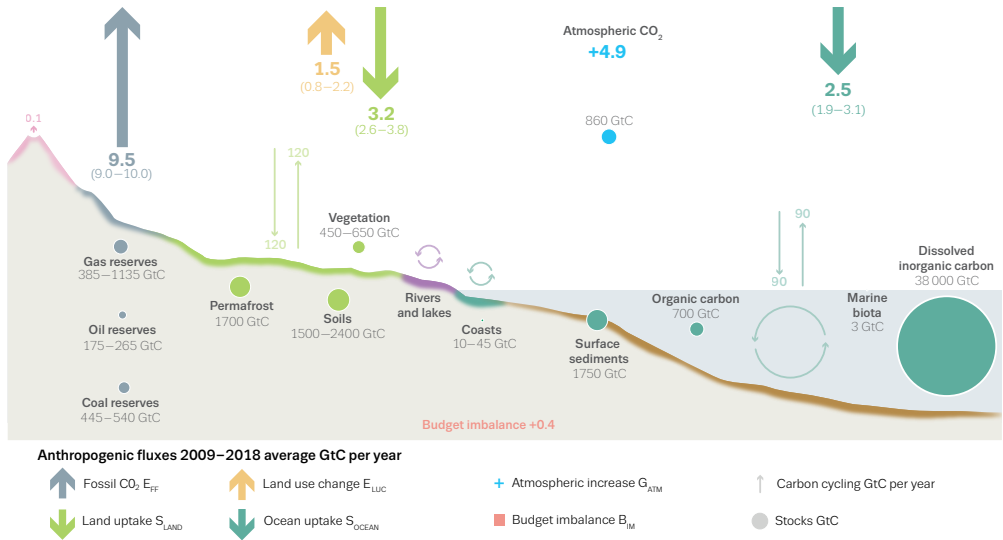


Figure 1.1. The global carbon cycle for the period 2009–2018, adapted from (Friedlingstein et al. 2019). The arrows represent fluxes between different carbon (C) pools (GtC·year⁻¹). The large arrows represent anthropogenic fluxes, while thin arrows represent the C cycle fluxes without anthropogenic influences. The C stock is given for each C pool (GtC).

biosphere (Friedlingstein et al. 2019) (Figure 1.1). The terrestrial biosphere thus acts as C sink, absorbing C that would otherwise remain in the atmosphere, and preventing even higher atmospheric CO₂ concentrations (Friedlingstein 2015). Forests, representing 80% of biomass globally (Bar-On et al. 2018) and storing 50–65% of terrestrial organic C (Reichstein and Carvalhais 2019), are an essential C sink estimated to be equal to 2.4 GtC y⁻¹ (Pan et al. 2011). The role of forests is thus fundamental in constraining future atmospheric CO₂ increase.

Projections of future climate and atmospheric CO₂ are created with numerical simulation models (Eyring et al. 2016). These projections are, however, uncertain. One of the sources of uncertainty in future climate is attributed to the magnitude of the terrestrial C sink (Ahlström et al. 2012; Ciais et al. 2013; Lovenduski and Bonan 2017; Schimel et al. 2015; Schurgers et al. 2018). This uncertainty results from the influence of contrasting factors among which higher CO₂ concentration, rising global temperature and N deposition (Huntzinger et al. 2017). While a rising CO₂ concentration is thought to drive photosynthesis and C uptake through the CO₂ fertilization effect (Ciais et al. 2013; Schimel et al. 2015), the influence of warming climate on these processes is less clear, with regional differences depending on the type of ecosystem (Ahlström et al. 2012; Huntzinger et al. 2017). One important source of the uncertainty in future terrestrial C uptake consists in the influence of N availability on vegetation growth (Ciais et al. 2013; Wieder et al. 2015). Accounting for N availability influences C sequestration and is

thought to decrease the magnitude of the CO₂ fertilization effect previously simulated when only accounting for rising CO₂ concentration (Huntzinger et al. 2017; Wieder et al. 2015).

1.3 N cycle and canopy N

N is an essential nutrient for plant growth being a major constituent of proteins and nucleic acids. The majority of the N found in leaves of C3 plants is associated with proteins of the photosynthetic apparatus. N is a constituent of chlorophyll and Rubisco, both of which are central to the photosynthetic process (Evans and Seemann 1989). Linear relationships exist between total foliar N content and both N content of Rubisco and leaf chlorophyll content (Evans 1989). Nitrogen has also been associated with physiological processes. At the leaf level, the leaf N concentration (%N) was found to be linked to photosynthetic capacity (Reich et al. 1999; Wright et al. 2004) and light use efficiency (Kergoat et al. 2008) across various environments and species.

At the site level too, canopy N (%N), defined as the leaf N scaled to the whole vegetation canopy (Smith and Martin 2001), was correlated to above-ground net primary productivity (NPP) (Reich 2012; Smith et al. 2002; Zhou et al. 2018) as well as canopy photosynthetic capacity (Ollinger et al. 2008). In the field, canopy N is measured by sampling multiple fresh leaves from all dominant species on the plot and analyzing their N concentration. The whole-stand canopy N concentration value is obtained by weighting the resulting species-specific leaf N values by the proportion of each species on the plot (Smith and Martin 2001).

Nitrogen is a limiting nutrient to plant growth in ecosystems worldwide. Low N availability results in physiological stress and reduced growth (Chapin 1987). Early N addition experiments showed that N addition to various N limited ecosystems led to an increase in primary productivity (LeBauer and Treseder 2008; Vitousek and Howarth 1991). Similarly, N fertilizers are also widely used in agriculture to produce higher yield. N fertilizers are produced by human fixation of N₂ from the atmosphere (Erisman et al. 2008). This process has influenced the N cycle to such an extent that the estimated annual amount of N₂ converted by humans from the atmosphere to reactive forms of N is 120 million tons, which exceeds the combined effects from all Earth's terrestrial processes (Rockström et al. 2009). Currently, approximately 50% of the global food production can be attributed to N fertilizers (Erisman et al. 2008). Intensive agriculture, among which the cultivation of leguminous crops, industry and traffic also contributed to this human induced enhancement of the N cycle (Erisman et al. 2011). Reactive forms of N cascade through the natural system and negatively affect the environment (Erisman et al. 2008; Galloway et al. 2008; Vitousek et al. 1997). One of the consequences of the enhancement of the N cycle is the increase in N deposition. Natural N deposition rates, i.e. without the anthropogenic influence, is estimated to be 0.5 kg N ha⁻¹ year⁻¹ at maximum (Galloway et al. 2008), while in some regions, in particular western Europe, current N deposition rates are up to 20 kg N ha⁻¹ year⁻¹ (Lu et al. 2013).

The constraint that N availability places on primary productivity as well as the human induced changes in N availability have implications for the C cycle. The link between the N and C cycles should thus be accounted for when studying the effects of increased atmospheric CO₂ on productivity (Erisman et al. 2011). The effect of elevated CO₂ concentration on C

sequestration in forest stands is studied using free-air CO₂ enrichment (FACE) experiments. FACE experiments allow to artificially raise the ambient CO₂ concentration to projected atmospheric CO₂ levels during extended periods of time as well as to manipulate N availability. The effect of N availability on the potential C sink is illustrated in two examples. In a sweetgum (*Liquidambar styraciflua*) forest stand with 550 ppm CO₂, the productivity was initially enhanced due to the CO₂ fertilization effect, but later declined due to low N availability (Norby et al. 2010). Also, in a loblolly pine (*Pinus taeda* L.) forest stand under elevated CO₂, no increase in wood C sequestration was observed before nutrient addition (Oren et al. 2001). Similarly, it was observed that nutrient-rich forests assimilate C more efficiently than nutrient-poor forest sites (Fernández-Martínez et al. 2014). Next to N, other nutrients, such phosphorous (P) are also limiting for vegetation growth (Chapin 1987) and P limitations influence primary productivity of the terrestrial biosphere (Elser et al. 2007; Vitousek et al. 2010).

At global scale, the general observation is that increased N deposition has enhanced C sequestration in forest ecosystems (Churkina et al. 2009; Fleischer et al. 2013; de Vries et al. 2014; De Vries et al. 2006). These results show that it is necessary to know better where and in which ecosystems N is an important factor when modelling C uptake by terrestrial vegetation (Fernández-Martínez et al. 2014). Forests, as a major C sink, are especially important (Reichstein and Carvalhais 2019). In this context, global vegetation models enable to quantify the current and future C pools and fluxes in the terrestrial system.

1.4 Vegetation models

Global vegetation models (GVMs) simulate the exchange of water, carbon and nutrients between vegetation and soil compartments by employing equations representing physiological processes (Zaehle and Friend 2010). In global carbon budget calculations, GVMs are used to model the terrestrial CO₂ sink as well as the emissions from land use, land use change and forestry (Friedlingstein et al. 2019). GVMs are also used to model the biosphere response to future changes in the C cycle and climate (Wärmlind et al. 2014). To account for the constraining effect of low N availability on vegetation growth, the N cycle is integrated in vegetation models (Smith et al. 2014; Zaehle and Friend 2010; Zaehle and Dalmonech 2011; Zaehle et al. 2014). In these models, the foliage N status of the plants is represented by either leaf N (Smith et al. 2014; Zaehle and Friend 2010) or canopy N (Ollinger and Smith 2005).

Using GVMs with C-N modules, the influence of N limitation on vegetation C uptake is simulated under future climate, CO₂ concentration and N deposition (Zaehle 2013; Zaehle et al. 2010). Models give somewhat conflicting results regarding projected C budgets. While the LPJ-GUESS GVM predicts an increase of the C sink due to N-deposition (Wärmlind et al. 2014), other models report a decrease of C assimilation up to 50% of the CO₂ fertilization effect following N limitation (Huntzinger et al. 2017). In another study, accounting for N limitation on the C storage lowered the projected future net primary production (NPP) by 19% (Wieder et al. 2015).

To untangle these different uncertainties, it is necessary to evaluate the models. Currently, coupled C-N vegetation models are evaluated based on their ability to properly simulate the C cycle. In order to evaluate model projections related to the N cycle module, data on the N

cycle are needed. However, compared to the C cycle, for which data are broadly available, observations on the N cycle are lacking at global scale (Zaehle and Dalmonech 2011). In this context, remote sensing could provide data on the spatial pattern of canopy N at large scale. Earth observation has already been used to evaluate C-cycle related simulations from GVMs (Exbrayat et al. 2019) and could similarly be used to evaluate the N-cycle components of GVMs. In chapter 5, the foliage N concentration output of two GVMs is compared with a canopy N map in European forests obtained in chapter 4. This map was created with a model built from data collected at forest plots, remote sensing data and environmental factors.

1.5 Canopy N remote sensing

Remote sensing of the biochemical concentration of elements in leaves, in particular N, evolved from near-infrared spectroscopy (NIRS) studying the relationship between dried and ground leaves concentration of N-containing organic compounds, mainly proteins, and their reflectance spectra (Curran 1989; Kokaly 2001; Kumar et al. 2006). The observed correlations between the biochemical concentration and the spectra were linked to absorption features caused by stretching frequencies of the N-H bonds present in the protein studied (Kokaly 2001; Kumar et al. 2006). This direct influence of N on the reflectance spectra is, however, not easily exploited for remote sensing of foliage N concentration due to multiple absorption features overlapping and interfering with each other as well as the strong absorption by water in the NIR region of the spectrum (Curran 1989; Kumar et al. 2006).

Leaf and canopy N remote sensing has instead been linked to chlorophyll absorption and the red-edge spectral region. The red-edge is a typical characteristic of vegetation spectral response located between 680 and 750 nm (Clevers et al. 2002; Curran et al. 1990). It is characterized by the sudden increase in reflectance between the low reflectance in the red region, due to chlorophyll absorption, and the high scattering observed in the near-infrared (NIR) region (Figure 1.2). The red-edge has been shown to be useful for chlorophyll detection in various environments (Curran et al. 1990; Dash and Curran 2004; Filella and Penuelas 1994). Owing to the strong correlation between foliar N and chlorophyll content (Evans 1989; Schlemmer et al. 2013, section 1.3), the red-edge has also been used for canopy N remote sensing (Clevers and Gitelson 2013; Clevers and Kooistra 2012; Fitzgerald et al. 2010; Li et al. 2014; Mutanga and Skidmore 2007; Ramoelo et al. 2012).

While initially NIRS studies focused on N content in dried agricultural products (Kumar et al. 2006), remote sensing of foliage was later extended to the study of fresh leaves (Curran et al. 1992). Since then, different applications of canopy N remote sensing have been developed. It is extensively used in agricultural studies as a tool to monitor N status from various crops, using either ground-based or airborne sensors, and enabling precise N fertilizer applications (Chen et al. 2010; Hansen and Schjoerring 2003; Muñoz-Huerta et al. 2013; Nigon et al. 2015; Xue et al. 2004; Zhu et al. 2007). From an ecological perspective, canopy N remote sensing has also been applied to study forest disturbances (Deel et al. 2012) as well as biological invasion (Asner and Vitousek 2005).

Canopy N remote sensing is based on empirical analysis methods. Stepwise multiple linear regression and partial least squares regression of remotely sensed information against

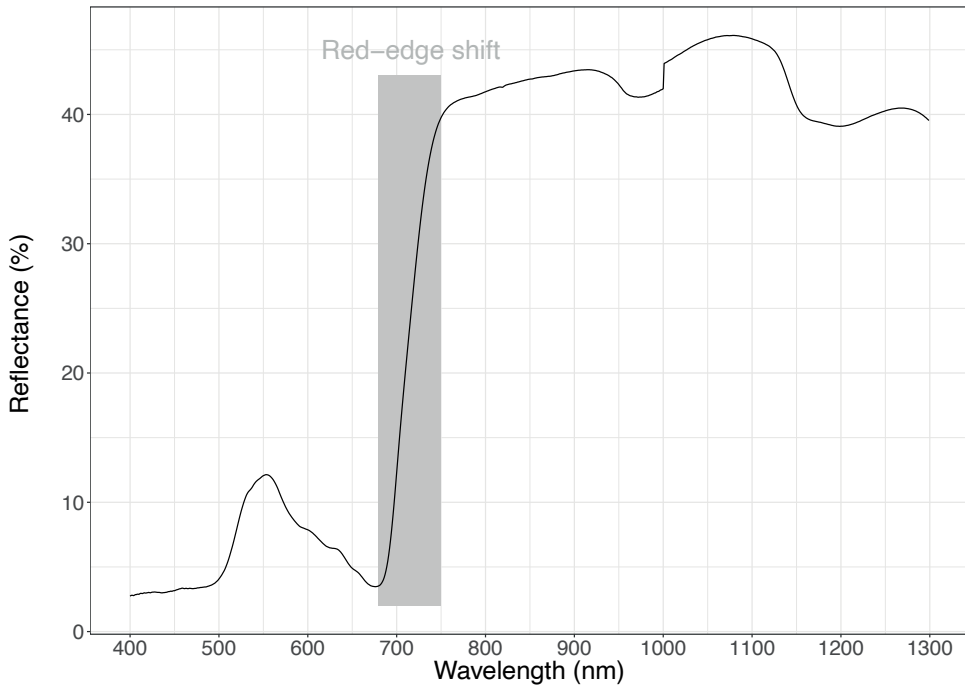


Figure 1.2. *Vegetation reflectance curve illustrating the red-edge shift between the red and NIR spectral regions.*

measurements of canopy N on the ground are frequently used (Berger et al. 2020; Kumar et al. 2006; Wang et al. 2016). Among other methods, vegetation indices (VIs) are one of the most straightforward methods to remotely sense canopy N. VIs are computed from several spectral bands most often combined in a ratio or normalized difference (Glenn et al. 2008). VIs do not rely on the full reflectance spectrum as it is possible to calculate VIs from a restricted number of bands, either broad or narrow, obtained from multispectral sensors (Thenkabail et al. 2002). Multiple VIs have been developed and are frequently used in agricultural studies where they are most often computed from handheld spectroradiometers measurements (Berger et al. 2020). Vegetation indices have also been applied in natural environments, e.g. in grasslands, and computed from either airborne (Ling et al. 2014; Mirik et al. 2005) or spaceborne (Mutanga et al. 2015; Ramoelo et al. 2012; Ullah et al. 2012) sensors at local scale. While less common, this method has also shown good results in forest environments using imaging spectrometry measurements at local scale and high spatial resolution from airborne (Serrano et al. 2002; Wang et al. 2016) or spaceborne sensors (Yu et al. 2017).

Since these results have been obtained in local scale studies, one of the main challenges when expanding canopy N remote sensing to larger areas is the laborious ground truth needed. Intensive field sampling campaigns are necessary over a large number of plots preferably occurring simultaneously with the sensor overflight. In an effort to expand canopy N remote sensing to larger areas, several studies combined canopy N sampling measurements from

multiple study sites (Martin et al. 2008; Ollinger et al. 2008; Singh et al. 2015), which increased the number of plots available to establish the statistical relationships. Other studies explore the possibility to use instruments with lower spatial and spectral resolution to map canopy N over larger areas (Lepine et al. 2016; Wallis et al. 2019).

Recently, a few studies explored the possibility to map foliar traits at large scale using machine learning methods (Campos-Taberner et al. 2018; Moreno-Martínez et al. 2018). These studies exploit already existing databases with foliar traits as well as long time series of multispectral sensors such as MODIS. This approach, which does not rely on simultaneous overflight of the sensor, enables to cover large areas as more plots are available to identify the statistical models. Machine learning methods, such as the random forests algorithm, have already been exploited to map canopy N at local scale in different ecosystems with promising results (Chemura et al. 2018; Mutowo et al. 2018; Ramoelo et al. 2015). The random forests algorithm is based on regressions trees and, compared to traditional linear models, allow to model non-linear relationships between numerous predictors and the target variable, with no assumption on the predictor's statistical distribution. This facilitates the inclusion of ancillary variables, such as climate, alongside earth observation data to build the model.

1.6 Research questions and outline of the thesis

Due to the role N plays as essential and limiting nutrient for plant growth (LeBauer and Treseder 2008; Vitousek and Howarth 1991), the N cycle influences and interacts with the C cycle. Climate change, caused by the increase in atmospheric CO₂ concentration (Bindoff et al. 2013; Friedlingstein et al. 2019), negatively affects life on earth and poses a threat to human communities (Mbow et al. 2019; Mirzabaev et al. 2019; Olsson et al. 2019). It is essential to understand how the terrestrial biosphere, which now acts as a C sink, will respond to increasing atmospheric CO₂ concentration caused by future CO₂ emissions (Friedlingstein et al. 2019; Lovenduski and Bonan 2017; Schurgers et al. 2018). Due to the link between the C and N cycles, the N cycle will influence the response of the terrestrial biosphere (Wieder et al. 2015). In order to understand and predict this response with GVMs, data are needed to train and evaluate the models. Data on the N cycle is however lacking at global scale. Applying remote sensing methods has the potential to bridge this gap in N data availability.

Spatially explicit information of canopy N is lacking over large areas. In this context, several scientific issues have not yet been addressed in current studies. Among the different remote sensing methods, it is not known whether VIs are suitable to estimate canopy N spatial patterns at large scale. Next to remote sensing, environmental variables could also be useful for canopy N mapping. Is it possible to integrate multiple data sources, both VIs and environmental variables, to examine the spatial pattern of canopy N at large scale? Finally, it is not known yet how the canopy N predictions from GVMs compare with canopy N map obtained from remote sensing.

In this thesis, the objective was to explore the feasibility of mapping canopy N over large areas using remote sensing and its application for ecosystem modelling studies. The thesis will address the following research questions which are further detailed below:

How well can we estimate canopy nitrogen (N) across spatial scales using vegetation indices (VIs) from remote sensing and environmental variables?

- (i) What is the accuracy of canopy N estimated from remotely sensed vegetation indices (VIs)? (**Chapter 2, 3 and 4**)
- (ii) How will canopy N estimation from remote sensing be influenced if environmental variables are included as predictors? (**Chapter 4**)
- (iii) How does canopy N estimated from remote sensing and environmental variables compare with foliage N simulated by global vegetation models (GVMs)? (**Chapter 5**)

To answer these questions, the next three chapters examine remote sensing of canopy N at different scales. In chapter 5, the canopy N map in European forests obtained in chapter 4 is compared to the output of two vegetation models.

In **chapter 2**, we performed an experiment in which we measured the spectral reflectance of a grass species (*Holcus lanatus L.*) with a spectroradiometer. The leaves were sampled to measure biochemical concentrations. The obtained spectra were used to compute VIs, both already existing and specifically designed for this dataset. The possibility of using VIs for canopy N remote sensing was tested by relating the VIs to the grasses canopy N as well as canopy P and canopy N:P ratio. The influence of the spectral resolution was evaluated by resampling the reflectance spectra to existing satellite spectral resolution. Results showed that canopy N:P was modest but significantly related to both existing vegetation indices ($r^2 = 0.16-0.48$) and optimized indices ($r^2 = 0.59-0.72$). Correlations were similar to what was observed for individual canopy N or canopy P.

In **chapter 3**, we related the MERIS Terrestrial Chlorophyll Index (MTCI) measured from the MERIS sensor aboard the Envisat satellite to canopy N concentration and canopy N content of Mediterranean forests in Catalonia, Spain. By doing so, we tested if it was possible to use a specific red-edge based VI for canopy N remote sensing at regional scale. The influence of the spatial resolution on the results was also explored. The relationship between MTCI and canopy N concentration was strongest for deciduous broadleaf and mixed plots and the relationship between MTCI and canopy N content was strongest for evergreen needleleaf trees.

In **chapter 4**, we spatially predicted canopy N in European forests using random forests models calibrated with canopy N data from the ICP Forest database. The influence of including environmental variables within the predictor variables was evaluated and proved valuable. Results showed that canopy N could be estimated both within and among forest types using the random forests technique and calibration data from ICP Forests with good accuracy.

In **chapter 5**, we compared the canopy N map obtained from the random forests model obtained in chapter 4 to spatial predictions of foliar N from two vegetation model, O-CN and LPJ-GUESS. The three models showed reasonable agreement in Europe regarding spatial patterns except for the Mediterranean area. When the lower latitude pixels were excluded, the three models showed significant linear relationships with each other. This showed that, while the physiological processes were represented differently in the two GVMs, the results obtained regarding canopy N are comparable and the N cycle representation in the models are coherent.

In **chapter 6**, we reflected and discussed on the implications of the results obtained in the previous chapters in regard of the scientific questions and objectives addressed in this thesis.

1.7 References

- Ahlström, A., Schurgers, G., Arneeth, A., & Smith, B. (2012). Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections. *Environmental Research Letters*, *7*, 044008. [10.1088/1748-9326/7/4/044008](https://doi.org/10.1088/1748-9326/7/4/044008)
- Allen, M.R., Dube, O.P., Solecki, W., Aragon-Durand, F., Cramer, W., Humphreys, S., Kainuma, M., Kala, J., Mahowald, N., Mulugetta, Y., Perez, R., Wairiu, M., & Zickfeld, K. (2018). Framing and Context. In V. Masson-Delmotte, P. Zhai, H.-O. Portner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Pean, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, & T.e. Waterfield (Eds.), *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* In Press.
- Asner, G.P., & Vitousek, P.M. (2005). Remote analysis of biological invasion and biogeochemical change. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 4383-4386. [10.1073/pnas.0500823102](https://doi.org/10.1073/pnas.0500823102)
- Bar-On, Y.M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences*, *115*, 6506. [10.1073/pnas.1711842115](https://doi.org/10.1073/pnas.1711842115)
- Berger, K., Verrelst, J., Féret, J.-B., Wang, Z., Woche, M., Strathmann, M., Danner, M., Mauser, W., & Hank, T. (2020). Crop nitrogen monitoring: Recent progress and principal developments in the context of imaging spectroscopy missions. *Remote Sensing of Environment*, *242*, 111758. <https://doi.org/10.1016/j.rse.2020.111758>
- Bindoff, N.L., Stott, P.A., AchutaRao, K.M., Allen, M.R., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S., Mokhov, I.I., Overland, J., Perlwitz, J., Sebbari, R., & Zhang, X. (2013). Detection and Attribution of Climate Change: from Global to Regional. In T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P.M.e. Midgley (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press
- Burrell, A.L., Evans, J.P., & De Kauwe, M.G. (2020). Anthropogenic climate change has driven over 5 million km² of drylands towards desertification. *Nature Communications*, *11*. [10.1038/s41467-020-17710-7](https://doi.org/10.1038/s41467-020-17710-7)
- Campos-Taberner, M., Moreno-Martínez, Á., García-Haro, J.F., Camps-Valls, G., Robinson, P.N., Kattge, J., & Running, W.S. (2018). Global Estimation of Biophysical Variables from Google Earth Engine Platform. *Remote Sensing*, *10*. [10.3390/rs10081167](https://doi.org/10.3390/rs10081167)
- Chapin, F.S.I. (1987). Adaptations and physiological responses of wild plants to nutrient stress. In W.H. Gabelman, & B.C. Loughman (Eds.), *Genetic Aspects of Plant Mineral Nutrition. Developments in Plant and Soil Sciences* Dordrecht: Springer. https://doi.org/10.1007/978-94-009-3581-5_2.
- Chemura, A., Mutanga, O., Odindi, J., & Kutuywayo, D. (2018). Mapping spatial variability of foliar nitrogen in coffee (*Coffea arabica* L.) plantations with multispectral Sentinel-2 MSI data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *138*, 1-11. <https://doi.org/10.1016/j.isprsjprs.2018.02.004>
- Chen, P., Haboudane, D., Tremblay, N., Wang, J., Vigneault, P., & Li, B. (2010). New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sensing of Environment*, *114*, 1987-1997. <https://doi.org/10.1016/j.rse.2010.04.006>

- Churkina, G., Brovkin, V., Von Bloh, W., Trusilova, K., Jung, M., & Dentener, F. (2009). Synergy of rising nitrogen depositions and atmospheric CO₂ on land carbon uptake moderately offsets global warming. *Global Biogeochemical Cycles*, 23. 10.1029/2008GB003291
- Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quere, C., Myneni, R.B., Piao, S., & Thornton, P. (2013). Carbon and other biogeochemical cycles. In T.F. Stocker, Q. D., G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P.M. Midgley (Eds.), *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 465-570). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press
- Clevers, J.G.P.W., De Jong, S.M., Epema, G.F., Van Der Meer, F.D., Bakker, W.H., Skidmore, A.K., & Scholte, K.H. (2002). Derivation of the red edge index using the MERIS standard band setting. *International Journal of Remote Sensing*, 23, 3169-3184. 10.1080/01431160110104647
- Clevers, J.G.P.W., & Gitelson, A.A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and-3. *International Journal of Applied Earth Observation and Geoinformation*, 23, 344-351. doi:10.1016/j.jag.2012.10.008
- Clevers, J.G.P.W., & Kooistra, L. (2012). Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5, 574-583. 10.1109/JSTARS.2011.2176468
- Cook, J., Nuccitelli, D., Green, S.A., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P., & Skuce, A. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters*, 8, 024024. 10.1088/1748-9326/8/2/024024
- Cook, J., Oreskes, N., Doran, P.T., Anderegg, W.R.L., Verheggen, B., Maibach, E.W., Carlton, J.S., Lewandowsky, S., Skuce, A.G., Green, S.A., Nuccitelli, D., Jacobs, P., Richardson, M., Winkler, B., Painting, R., & Rice, K. (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters*, 11, 048002. 10.1088/1748-9326/11/4/048002
- Curran, P.J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*, 30, 271-278. [https://doi.org/10.1016/0034-4257\(89\)90069-2](https://doi.org/10.1016/0034-4257(89)90069-2)
- Curran, P.J., Dungan, J.L., & Gholz, H.L. (1990). Exploring the relationship between reflectance red edge and chlorophyll content in slash pine. *Tree Physiology*, 7, 33-48
- Curran, P.J., Dungan, J.L., Macler, B.A., Plummer, S.E., & Peterson, D.L. (1992). Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. *Remote Sensing of Environment*, 39, 153-166. [https://doi.org/10.1016/0034-4257\(92\)90133-5](https://doi.org/10.1016/0034-4257(92)90133-5)
- Dash, J., & Curran, P.J. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25, 5403-5413. doi:10.1080/0143116042000274015
- de Vries, W., Du, E., & Butterbach-Bahl, K. (2014). Short and long-term impacts of nitrogen deposition on carbon sequestration by forest ecosystems. *Current Opinion in Environmental Sustainability*, 9-10, 90-104. <https://doi.org/10.1016/j.cosust.2014.09.001>
- De Vries, W.I.M., Reinds, G.J., Gundersen, P.E.R., & Sterba, H. (2006). The impact of nitrogen deposition on carbon sequestration in European forests and forest soils. *Global Change Biology*, 12, 1151-1173. 10.1111/j.1365-2486.2006.01151.x
- Deel, L.N., McNeil, B.E., Curtis, P.G., Serbin, S.P., Singh, A., Eshleman, K.N., & Townsend, P.A. (2012). Relationship of a Landsat cumulative disturbance index to canopy nitrogen and forest structure. *Remote Sensing of Environment*, 118, 40-49. 10.1016/j.rse.2011.10.026

- Elser, J.J., Bracken, M.E.S., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., Ngai, J.T., Seabloom, E.W., Shurin, J.B., & Smith, J.E. (2007). Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters*, *10*, 1135-1142. <https://doi.org/10.1111/j.1461-0248.2007.01113.x>
- Erismann, J.W., Galloway, J., Seitzinger, S., Bleeker, A., & Butterbach-Bahl, K. (2011). Reactive nitrogen in the environment and its effect on climate change. *Current Opinion in Environmental Sustainability*, *3*, 281-290. 10.1016/j.cosust.2011.08.012
- Erismann, J.W., Sutton, M.A., Galloway, J., Klimont, Z., & Winiwarter, W. (2008). How a century of ammonia synthesis changed the world. *Nature Geoscience*, *1*, 636-639. 10.1038/ngeo325
- Evans, J.R. (1989). Photosynthesis and nitrogen relationships in leaves of C3 plants. *Oecologia*, *78*, 9-19. doi:10.1007/BF00377192
- Evans, J.R., & Seemann, J.R. (1989). The allocation of protein nitrogen in the photosynthetic apparatus: costs, consequences, and control. In W.R. Briggs (Ed.), *Photosynthesis* New York: Alan R. Liss, Inc.
- Exbrayat, J.F., Bloom, A.A., Carvalhais, N., Fischer, R., Huth, A., MacBean, N., & Williams, M. (2019). Understanding the Land Carbon Cycle with Space Data: Current Status and Prospects. *Surveys in Geophysics*, *40*, 735-755. 10.1007/s10712-019-09506-2
- Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., & Taylor, K.E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, *9*, 1937-1958. 10.5194/gmd-9-1937-2016
- Fernández-Martínez, M., Vicca, S., Janssens, I.A., Sardans, J., Luysaert, S., Campioli, M., Chapin Iii, F.S., Ciais, P., Malhi, Y., Obersteiner, M., Papale, D., Piao, S.L., Reichstein, M., Rodà, F., & Peñuelas, J. (2014). Nutrient availability as the key regulator of global forest carbon balance. *Nature Climate Change*, *4*, 471-476. 10.1038/nclimate2177
- Filella, I., & Penuelas, J. (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *International Journal of Remote Sensing*, *15*, 1459-1470. <http://dx.doi.org/10.1080/01431169408954177>
- Fitzgerald, G., Rodriguez, D., & O'Leary, G. (2010). Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index – The canopy chlorophyll content index (CCCI). *Field Crops Research*, *116*, 318-324. <http://dx.doi.org/10.1016/j.fcr.2010.01.010>
- Fleischer, K., Rebel, K.T., Van Der Molen, M.K., Erismann, J.W., Wassen, M.J., Van Loon, E.E., Montagnani, L., Gough, C.M., Herbst, M., Janssens, I.A., Gianelle, D., & Dolman, A.J. (2013). The contribution of nitrogen deposition to the photosynthetic capacity of forests. *Global Biogeochemical Cycles*, *27*, 187-199. 10.1002/gbc.20026
- Friedlingstein, P. (2015). Carbon cycle feedbacks and future climate change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*. <https://doi.org/10.1098/rsta.2014.0421>
- Friedlingstein, P., Jones, M.W., O'Sullivan, M., Andrew, R.M., Hauck, J., Peters, G.P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D.C.E., Canadell, J.G., Ciais, P., Jackson, R.B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L.P., Currie, K.I., Feely, R.A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D.S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R.A., Hurtt, G., Ilyina, T., Jain, A.K., Joetzjer, E., Kaplan, J.O., Kato, E., Klein Goldewijk, K., Korsbakken, J.I., Landschützer, P., Lauvset, S.K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P.C., Melton, J.R., Metzl, N.,

- Munro, D.R., Nabel, J.E.M.S., Nakaoka, S.I., Neill, C., Omar, A.M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P.P., Tian, H., Tilbrook, B., Tubiello, F.N., van der Werf, G.R., Wiltshire, A.J., & Zaehle, S. (2019). Global Carbon Budget 2019. *Earth Syst. Sci. Data*, *11*, 1783-1838. 10.5194/essd-11-1783-2019
- Galloway, J.N., Townsend, A.R., Erisman, J.W., Bekunda, M., Cai, Z., Freney, J.R., Martinelli, L.A., Seitzinger, S.P., & Sutton, M.A. (2008). Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions. *Science*, *320*, 889. 10.1126/science.1136674
- Glenn, E.P., Huete, A.R., Nagler, P.L., & Nelson, S.G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*, *8*, 2136-2160. doi:10.3390/s8042136
- Hansen, P.M., & Schjoerring, J.K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, *86*, 542-553. doi:10.1016/S0034-4257(03)00131-7
- Huntzinger, D.N., Michalak, A.M., Schwalm, C., Ciais, P., King, A.W., Fang, Y., Schaefer, K., Wei, Y., Cook, R.B., Fisher, J.B., Hayes, D., Huang, M., Ito, A., Jain, A.K., Lei, H., Lu, C., Maignan, F., Mao, J., Parazoo, N., Peng, S., Poulter, B., Ricciuto, D., Shi, X., Tian, H., Wang, W., Zeng, N., & Zhao, F. (2017). Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Scientific Reports*, *7*. 10.1038/s41598-017-03818-2
- Jia, G., Shevliakova, E., Artaxo, P., De Noblet-Ducoudré, N., Houghton, R., House, J., Kitajima, K., Lennard, C., Popp, A., Sirin, A., Sukumar, R., & Vercho, L. (2019). Land-climate interactions. In P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, & J.e. Malley (Eds.), *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems* In press.
- Kergoat, L., Lafont, S., Arneth, A., Le Dantec, V., & Saugier, B. (2008). Nitrogen controls plant canopy light-use efficiency in temperate and boreal ecosystems. *Journal of Geophysical Research: Biogeosciences*, *113*. doi:10.1029/2007JG000676
- Kokaly, R.F. (2001). Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment*, *75*, 153-161. 10.1016/S0034-4257(00)00163-2
- Kumar, L., Schmidt, K., Dury, S., & Skidmore, A. (2006). Imaging Spectrometry and Vegetation Science. In F.D.v.d. Meer, & S.M. de Jong (Eds.), *Imaging Spectrometry: Basic Principles and Prospective Applications* (pp. 111-155). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-0-306-47578-8_5. ISBN: 978-0-306-47578-8
- LeBauer, D.S., & Treseder, K.K. (2008). Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, *89*, 371-379. 10.1890/06-2057.1
- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., & Martin, M.E. (2016). Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sensing of Environment*, *173*, 174-186. doi:10.1016/j.rse.2015.11.028
- Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., Liu, Y., Liu, B., Ustin, S.L., & Chen, X. (2014). Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Research*, *157*, 111-123. doi:10.1016/j.fcr.2013.12.018

- Ling, B., Goodin, D.G., Mohler, R.L., Laws, A.N., & Joern, A. (2014). Estimating canopy nitrogen content in a heterogeneous grassland with varying fire and grazing treatments: Konza Prairie, Kansas, USA. *Remote Sensing*, 6, 4430-4453. 10.3390/rs6054430
- Lovenduski, N.S., & Bonan, G.B. (2017). Reducing uncertainty in projections of terrestrial carbon uptake. *Environmental Research Letters*, 12. 10.1088/1748-9326/aa66b8
- Lu, X., Jiang, H., Zhang, X., Liu, J., Zhang, Z., Jin, J., Wang, Y., Xu, J., & Cheng, M. (2013). Estimated global nitrogen deposition using NO₂ column density. *International Journal of Remote Sensing*, 34, 8893-8906. 10.1080/01431161.2013.853894
- Martin, M.E., Plourde, L.C., Ollinger, S.V., Smith, M.L., & McNeil, B.E. (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment*, 112, 3511-3519. doi:10.1016/j.rse.2008.04.008
- Mbow, C., Rosenzweig, C., Barioni, L.G., Benton, T.G., Herrero, M., Krishnapillai, M., Liwenga, E., Pradhan, P., Rivera-Ferre, M.G., Sapkota, T., Tubiello, F.N., & Xu, Y. (2019). Food Security. In P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, & J.e. Malley (Eds.), *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems* In press.
- 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. doi:10.2111/04-17.1
- Mirzabaev, A., Wu, J., Evans, J., García-Oliva, F., Hussein, I.A.G., Iqbal, M.H., Kimutai, J., Knowles, T., Meza, F., Nedjraoui, D., Tena, F., Türkeş, M., Vázquez, R.J., & We, M. (2019). Desertification. In P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, & J.e. Malley (Eds.), *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems* In press.
- Moreno-Martínez, Á., Camps-Valls, G., Kattge, J., Robinson, N., Reichstein, M., van Bodegom, P., Kramer, K., Cornelissen, J.H.C., Reich, P., Bahn, M., Niinemets, Ü., Peñuelas, J., Craine, J.M., Cerabolini, B.E.L., Minden, V., Laughlin, D.C., Sack, L., Allred, B., Baraloto, C., Byun, C., Soudzilovskaia, N.A., & Running, S.W. (2018). A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sensing of Environment*, 218, 69-88. <https://doi.org/10.1016/j.rse.2018.09.006>
- Muñoz-Huerta, R.F., Guevara-Gonzalez, R.G., Contreras-Medina, L.M., Torres-Pacheco, I., Prado-Olivarez, J., & Ocampo-Velazquez, R.V. (2013). A review of methods for sensing the nitrogen status in plants: Advantages, disadvantages and recent advances. *Sensors (Switzerland)*, 13, 10823-10843. 10.3390/s130810823
- Mutanga, O., Adam, E., Adjorloloa, C., & Abdel-Rahmanw, E.M. (2015). Evaluating the robustness of models developed from field spectral data in predicting African grass foliar nitrogen concentration using WorldView-2 image as an independent test dataset. *International Journal of Applied Earth Observation and Geoinformation*, 34, 178-187. 10.1016/j.jag.2014.08.008
- 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. 10.1016/j.isprsjprs.2007.02.001

- Mutowo, G., Mutanga, O., & Masocha, M. (2018). Evaluating the Applications of the Near-Infrared Region in Mapping Foliar N in the Miombo Woodlands. *Remote Sensing*, *10*. 10.3390/rs10040505
- Nigon, T.J., Mulla, D.J., Rosen, C.J., Cohen, Y., Alchanatis, V., Knight, J., & Rud, R. (2015). Hyperspectral aerial imagery for detecting nitrogen stress in two potato cultivars. *Computers and Electronics in Agriculture*. 10.1016/j.compag.2014.12.018
- Norby, R.J., Warren, J.M., Iversen, C.M., Medlyn, B.E., & McMurtrie, R.E. (2010). CO₂ enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences*, *107*, 19368. 10.1073/pnas.1006463107
- Ollinger, S.V., Richardson, A.D., Martin, M.E., Hollinger, D.Y., Frohling, S.E., Reich, P.B., Plourde, L.C., Katul, G.G., Munger, J.W., Oren, R., Smith, M.L., Paw U, K.T., Bolsta, P.V., Cook, B.D., Day, M.C., Martin, T.A., Monson, R.K., & 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 Sciences of the United States of America*, *105*, 19336-19341. doi:10.1073/pnas.0810021105.
- Ollinger, S.V., & 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*, *8*, 760-778. doi:10.1007/s10021-005-0079-5
- Olsson, L., H. Barbosa, H., Bhadwal, S., Cowie, A., Delusca, K., Flores-Renteria, D., Hermans, K., Jobbagy, E., Kurz, W., Li, D., Sonwa, D.J., & Stringer, L. (2019). Land Degradation. In P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, & J.e. Malley (Eds.), *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems* In press.
- Oren, R., Ellsworth, D.S., Johnsen, K.H., Phillips, N., Ewers, B.E., Maier, C., Schäfer, K.V.R., McCarthy, H., Hendrey, G., McNulty, S.G., & Katul, G.G. (2001). Soil fertility limits carbon sequestration by forest ecosystems in a CO₂-enriched atmosphere. *Nature*, *411*, 469-472. 10.1038/35078064
- Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., & Hayes, D. (2011). A Large and Persistent Carbon Sink in the World's Forests. *Science*, *333*, 988. 10.1126/science.1201609
- Ramoelo, A., Cho, M.A., Mathieu, R., Madonsela, S., van de Kerchove, R., Kaszta, Z., & Wolff, E. (2015). Monitoring grass nutrients and biomass as indicators of rangeland quality and quantity using random forest modelling and WorldView-2 data. *International Journal of Applied Earth Observation and Geoinformation*, *43*, 43-54. <https://doi.org/10.1016/j.jag.2014.12.010>
- Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., & Heitkönig, I.M.A. (2012). Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor. *International Journal of Applied Earth Observation and Geoinformation*, *19*, 151-162. doi:10.1016/j.jag.2012.05.009
- Reich, P.B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 2128-2134. doi:10.1098/rspb.2011.2270
- Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Volin, J.C., & Bowman, W.D. (1999). Generality of leaf trait relationships: A test across six biomes. *Ecology*, *80*, 1955-1969. doi:10.2307/176671

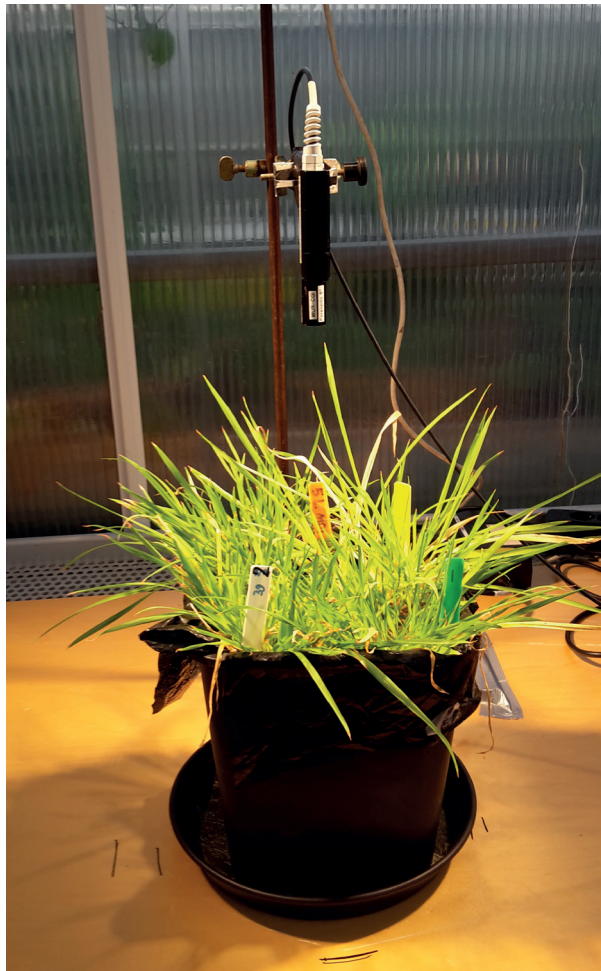
- Reichstein, M., & Carvalhais, N. (2019). Aspects of Forest Biomass in the Earth System: Its Role and Major Unknowns. *Surveys in Geophysics*, *40*, 693-707. 10.1007/s10712-019-09551-x
- Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin Iii, F.S., Lambin, E., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörilin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., & Foley, J. (2009). Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society*, *14*. 10.5751/ES-03180-140232
- Schimmel, D., Stephens, B.B., & Fisher, J.B. (2015). Effect of increasing CO₂ on the terrestrial carbon cycle. *Proceedings of the National Academy of Sciences*, *112*, 436. 10.1073/pnas.1407302112
- Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, *25*, 47-54. doi:10.1016/j.jag.2013.04.003
- Schurgers, G., Ahlström, A., Arneth, A., Pugh, T.A.M., & Smith, B. (2018). Climate Sensitivity Controls Uncertainty in Future Terrestrial Carbon Sink. *Geophysical Research Letters*, *45*, 4329-4336. 10.1029/2018GL077528
- Serrano, L., Peñuelas, J., & Ustin, S.L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*, 355-364. doi:10.1016/S0034-4257(02)00011-1
- Singh, A., Serbin, S.P., McNeil, B.E., Kingdon, C.C., & Townsend, P.A. (2015). Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications*, *25*, 2180-2197. 10.1890/14-2098.1
- Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, *11*, 2027-2054. doi:10.5194/bg-11-2027-2014
- Smith, M.L., & Martin, M.E. (2001). A plot-based method for rapid estimation of forest canopy chemistry. *Canadian Journal of Forest Research*, *31*, 549-555. doi:10.1139/x00-187
- 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. *Ecological Applications*, *12*, 1286-1302. doi:10.2307/3099972
- Thenkabail, P.S., Smith, R.B., & De Pauw, E. (2002). Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. *Photogrammetric Engineering and Remote Sensing*, *68*, 607-621
- Ullah, S., Si, Y., Schlerf, M., Skidmore, A.K., Shafique, M., & Iqbal, I.A. (2012). Estimation of grassland biomass and nitrogen using MERIS data. *International Journal of Applied Earth Observation and Geoinformation*, *19*, 196-204. doi:10.1016/j.jag.2012.05.008
- United Nations (2020). *The sustainable Development Goals Report 2020*. <https://sdgs.un.org/publications/sustainable-development-goals-report-2020-24686> ISBN: 978-92-1-101425-9
- Vitousek, P.M., Aber, J.D., Howarth, R.W., Likens, G.E., Matson, P.A., Schindler, D.W., Schlesinger, W.H., & Tilman, D. (1997). Human alteration of the global nitrogen cycle: Sources and consequences. *Ecological Applications*, *7*, 737-750. 10.2307/2269431
- Vitousek, P.M., & Howarth, R.W. (1991). Nitrogen limitation on land and in the sea: How can it occur? *Biogeochemistry*, *13*, 87-115. 10.1007/BF00002772

- Vitousek, P.M., Porder, S., Houlton, B.Z., & Chadwick, O.A. (2010). Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen-phosphorus interactions. *Ecological Applications*, 20, 5-15. <https://doi.org/10.1890/08-0127.1>
- Wallis, C.I.B., Homeier, J., Peña, J., Brandl, R., Farwig, N., & Bendix, J. (2019). Modeling tropical montane forest biomass, productivity and canopy traits with multispectral remote sensing data. *Remote Sensing of Environment*, 225, 77-92. <https://doi.org/10.1016/j.rse.2019.02.021>
- Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A.K., Jones, S., Suarez, L., Woodgate, W., Heiden, U., Heurich, M., & Hearne, J. (2016). Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sensing*, 8. doi:10.3390/rs8060491
- Wärmland, D., Smith, B., Hickler, T., & Arneith, A. (2014). Nitrogen feedbacks increase future terrestrial ecosystem carbon uptake in an individual-based dynamic vegetation model. *Biogeosciences*, 11, 6131-6146. 10.5194/bg-11-6131-2014
- Wieder, W.R., Cleveland, C.C., Smith, W.K., & Todd-Brown, K. (2015). Future productivity and carbon storage limited by terrestrial nutrient availability. *Nature Geoscience*, 8, 441. 10.1038/ngeo2413
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., & Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, 428, 821-827. doi:10.1038/nature02403
- Xue, L., Cao, W., Luo, W., Dai, T., & Zhu, Y. (2004). Monitoring Leaf Nitrogen Status in Rice with Canopy Spectral Reflectance. *Agronomy Journal*, 96, 135-142. 10.2134/agronj2004.0135
- Yu, Q., Wang, S., & Zhou, L. (2017). Investigation into the role of canopy structure traits and plant functional types in modulating the correlation between canopy nitrogen and reflectance in a temperate forest in northeast China. *Journal of Applied Remote Sensing*, 11. 10.1117/1.JRS.11.046013
- Zaehle, S. (2013). Terrestrial nitrogen-carbon cycle interactions at the global scale. *Philosophical Transactions of the Royal Society B*, 368. <http://doi.org/10.1098/rstb.2013.0125>
- Zaehle, S., & Dalmonech, D. (2011). Carbon-nitrogen interactions on land at global scales: Current understanding in modelling climate biosphere feedbacks. *Current Opinion in Environmental Sustainability*, 3, 311-320. 10.1016/j.cosust.2011.08.008
- Zaehle, S., Friedlingstein, P., & Friend, A.D. (2010). Terrestrial nitrogen feedbacks may accelerate future climate change. *Geophysical Research Letters*, 37. <https://doi.org/10.1029/2009GL041345>
- Zaehle, S., & Friend, A.D. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, 24. 10.1029/2009GB003521
- Zaehle, S., Medlyn, B.E., De Kauwe, M.G., Walker, A.P., Dietze, M.C., Hickler, T., Luo, Y., Wang, Y.-P., El-Masri, B., Thornton, P., Jain, A., Wang, S., Warland, D., Weng, E., Parton, W., Iversen, C.M., Gallet-Budynek, A., McCarthy, H., Finzi, A., Hanson, P.J., Prentice, I.C., Oren, R., & Norby, R.J. (2014). Evaluation of 11 terrestrial carbon-nitrogen cycle models against observations from two temperate Free-Air CO₂ Enrichment studies. *New Phytologist*, 202, 803-822. 10.1111/nph.12697
- Zhou, Z., Ollinger, S.V., & Lepine, L. (2018). Landscape variation in canopy nitrogen and carbon assimilation in a temperate mixed forest. *Oecologia*, 188, 595-606. 10.1007/s00442-018-4223-2

Zhu, Y., Tian, Y., Yao, X., Liu, X., & Cao, W. (2007). Analysis of common canopy reflectance spectra for indicating leaf nitrogen concentrations in wheat and rice. *Plant Production Science*, *10*, 400-411. 10.1626/pp.s.10.400

Chapter 2

Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses



Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses¹

Abstract. Reduced availability of plant nutrients such as nitrogen (N) and phosphorous (P) has detrimental effects on plant growth. Plant N:P ratio, calculated as the quotient of N and P concentrations, is an ecological indicator of relative N and P limitation. Remote sensing has already been widely used to detect plant traits in foliage, particularly canopy N and P concentrations and could be used to detect canopy N:P faster and at lower cost than traditional destructive methods. Despite the potential opportunity of applying remote sensing techniques to detect canopy N:P, studies investigating canopy N:P remote detection are scarce. In this study, we examined if vegetation indices developed for canopy N or P detection can also be used for canopy N:P detection. Using in situ spectrometry, we measured the reflectance of a common grass species, Yorkshire fog (*Holcus lanatus* L.), grown under different nutrient ratios and levels. We calculated 60 VIs found in literature and compared them to optimized VIs developed specifically for this study. The VIs were calculated using both the original narrow band spectra and the spectra resampled to the band properties of six satellite sensors (MSI-Sentinel 2, OLCI-Sentinel 3, MODIS-Terra/Aqua, OLI-Landsat 8, WorldView 4 and RapidEye) to investigate the influence of bandwidths and band positions. The results showed that canopy N:P was significantly related to both existing VIs ($r^2 = 0.16-0.48$) and optimized VIs ($r^2 = 0.59-0.72$) with correlations similar to what was observed for canopy N or canopy P. Existing VIs calculated with MSI and OLI sensors bands showed higher correlation with canopy N:P compared to the other sensors while the correlation with optimized VIs was not affected by the differences in sensors' bands. This study might lead to future practical applications using in situ reflectance measurements to sense canopy N:P in grasslands.

Keywords: Nutrient limitation, Canopy N:P, *Holcus lanatus* L., Remote sensing, Spectroradiometer, Satellite sensors

1 This chapter is based on: Loozen, Y., Karssenbergh, D., de Jong, S.M., Wang, S., van Dijk, J., Wassen, M.J., & Rebel, K.T. (2019). Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses. *International Journal of Applied Earth Observation and Geoinformation*, 75, 1-14. <https://doi.org/10.1016/j.jag.2018.08.012>

2.1 Introduction

Nutrients play an essential role in plant growth and foliage nutrient concentration is linked to several physiological and ecosystem processes. Chlorophyll content, photosynthetic capacity, leaf life span, light use efficiency and biomass primary productivity have all been related to foliar nitrogen (N) concentration at the leaf or canopy level (Bakker et al. 2011; Evans 1989; Green et al. 2003; Kergoat et al. 2008; Reich 2012; Reich et al. 1999; Wright et al. 2004). Similarly, foliage phosphorus (P) has been correlated to leaf life span and photosynthetic capacity (Wright et al. 2004) and influences the photosynthetic rate (V_{cmax}) (Walker et al. 2014). Non-optimal levels of foliar N and P will often result in reduced plant growth (LeBauer and Treseder 2008; Vitousek and Howarth 1991).

Plant N:P ratio, defined as the quotient of plant N and plant P concentrations and expressed in g N/g P (Güsewell 2004), is an ecological indicator of the relative N and P limitation (Koerselman and Meuleman 1996; Olde Venterink et al. 2003; Wassen et al. 1995). Several studies have identified threshold plant N:P ratio values for N and P limitation in different ecosystems (Güsewell et al. 2003; Koerselman and Meuleman 1996; Li et al. 2011; Olde Venterink et al. 2003; Tessier and Raynal 2003). Although threshold values defined by different authors may vary, in general it is safe to conclude that plant N:P ratio values lower than 10 tend to indicate N limitation, while values higher than 20 are an indication of P deficiency (Güsewell 2004). However, this ratio should be understood as a continuous gradient including a range of values (c. 10 to c. 20) in which co-limitation may occur where both N and P are in low supply (Güsewell 2004). At vegetation level, the N:P ratio is an informative variable that not only indicates potential limitation of N and P but which is also related to species composition, species richness, productivity and functional trait composition (Fujita et al. 2014; Roeling et al. 2018; Wassen et al. 2005).

Foliar N and P concentration can be measured in the laboratory, but this is labor intensive and expensive, especially for large sample sizes. An alternative approach consists of using remote sensing to estimate plant traits, among which foliage nutrient concentration (Homolová et al. 2013). Foliage N concentration influences the reflectance spectra through specific absorption features attributed to N-bonds and protein absorption. These absorption features are located at 1020 nm, 1510 nm, 1940 nm, 2060 nm, 2180 nm, 2300 nm and 2350 nm (Kumar et al. 2006). However, due to the overlapping absorption features of multiple compounds as well as the strong absorption by water in the shortwave-infrared region (SWIR, 1400-3000 nm), the interpretation of the reflectance spectra is difficult (Kumar et al. 2006). For this reason, foliage N detection by remote sensing is mainly linked to chlorophyll absorption and often includes regions of the spectrum associated with chlorophyll detection, i.e. the red-edge and near-infrared (NIR) regions (Schlemmer et al. 2013). Field spectrometry has been extensively applied to estimate canopy N in a variety of crops using so called spectral vegetation indices (VIs), which consist of a combination of spectral reflectance bands (Hansen and Schjoerring 2003; Li et al. 2014; Schlemmer et al. 2013; Tian et al. 2011). Other ecosystems have been investigated for canopy N estimation using VIs computed from airborne or spaceborne sensors at different spatial resolutions, including temperate forests (HySpex airborne sensor at 3 m, Wang et al. (2016b)), tropical forests (RapidEye satellite sensor at 5m, Cho et al. (2013)), Mediterranean forests (MERIS satellite sensor at 1 km, Loozen et al. (2018)),

as well as chaparral vegetation (AVIRIS airborne sensor at 18 m, Serrano et al. (2002)) and savannah (RapidEye satellite sensor at 5m, Ramoelo et al. (2012)).

Similar to canopy N, canopy P has been estimated using spectral indices, though to a lesser extent. Several studies have aimed to develop VIs for canopy P estimation in agricultural lands (Kawamura et al. 2011; Mahajan et al. 2014; Pimstein et al. 2011) and in a *Carex* dominated grassland (Wang et al. 2016a), all using field spectrometry.

VIs calculated from spectrometry measurements have already been extensively studied for canopy N detection (Pacheco-Labrador et al. 2014) and high accuracy results have been obtained (Tian et al. 2011). Different categories of VIs exist, the most common ones being the simple ratio (SR), the normalized difference (ND) and the simple difference (SD) (Tian et al. 2011; Wang et al. 2016a). A VI can be classified as either narrowband or broadband (Thenkabail et al. 2002). Narrowband VIs are computed from narrow reflectance bands measured through imaging spectrometry, which includes in situ reflectance measurement using a spectroradiometer as well as reflectance measured by specific airborne or spaceborne sensors (Tian et al. 2011; Wang et al. 2015). Broadband VIs are traditionally obtained from multispectral satellite sensors although they can also be computed from resampled narrow reflectance bands obtained from other sources (Clevers and Gitelson 2013). A common approach to develop VIs is by selecting the optimal indices from all pair-wise combinations of wavelengths in the visible, NIR and shortwave-infrared (SWIR) regions (400-2500 nm), i.e. a band combination analysis (Hansen and Schjoerring 2003; Tian et al. 2011).

Canopy N:P is relevant to ecological studies as it is a nutrient limitation indicator. It is thus somewhat surprising that remote sensing of canopy N:P using VIs has rarely been investigated. One study investigated canopy N:P detection in savanna grasslands using field spectrometry coupled with partial least squares regression with significant results ($r^2 = 0.69-0.85$, (Ramoelo et al. 2013)). Canopy N:P detection was also investigated in a boreal forest using vegetation indices calculated from airborne (Compact Airborne Spectrographic Imager, CASI) and spaceborne (Hyperion EO-1) imaging spectrometry (Gökkaya et al. 2015). The results showed that the VIs could be related to canopy N:P with r^2 ranging from 0.34 to 0.70. In this context, there is a need to examine in more detail how spectral properties, either from in situ measurement or satellite sensors, influence the performance of VIs regarding canopy N:P detection.

In this study, we aim at identifying if existing VIs that have been used for the estimation of canopy N or canopy P can also be used to remotely sense canopy N:P. To do so, we will evaluate three sub questions. How do VIs perform for canopy N:P estimation compared to canopy N and canopy P estimation? How do existing VIs perform compared to optimized VIs developed using the data collected in this study? How is the correlation between canopy N:P and the VIs influenced by different sensors' bandwidths and band positions?

To create a dataset representing a large range of canopy N:P values, we chose to execute the study under controlled conditions. We have grown a common temperate grassland species, *Holcus lanatus* L. using six nutrient treatments that reflect N limited, P limited, and N and P co-limited conditions, under high and low nutrient availability, respectively. Existing VIs found in the literature were computed using the original narrow band reflectance spectra measured by in situ spectrometry as well as the resampled spectra corresponding to six different satellites sensors (MSI aboard Sentinel 2, OLCI aboard Sentinel 3, MODIS aboard Terra-Aqua, OLI

aboard Landsat 8, WorldView 4 and RapidEye). We compared the existing VIs to optimized VIs, obtained using a band combination analysis carried out on both the original narrow band spectra and the resampled broadband spectra.

2.2 Material and methods

2.2.1 Culture of the plants

Holcus lanatus L. (Yorkshire fog) is a perennial common grass species, generally found on nutrient-rich soils throughout Europe. Seeds were collected from the Middenduin nature reserve (52 °24'N 4 °35'E) located in the western Netherlands and cultivated under controlled conditions in a greenhouse of Utrecht University. Seeds were sown on moist quartz sand and were transplanted after 22 days into pots containing a mixture of quartz sand (Carlo Bernasconi, Zürich, CH, 0.1-0.7 mm) and sand collected close to the area where seed were harvested under a ratio of 11:1. This was done to introduce soil fauna, bacteria and fungi for a complete soil community. Each pot was planted with four seedlings. Plants were grown further under nutrient treatments that lasted from July 2015 until June 2016 in the greenhouse with 400 Wm⁻² light from 9:00 am to 4:00 pm and a temperature regime that mimicked the temperate conditions of the Netherlands. Six nutrient treatments were applied with three N:P ratios, each in high and low nutrient supply levels. The nutrient ratios were N:P = 5 (N-limitation), N:P = 15 (balanced supply) and N:P = 45 (P-limitation). Each nutrient treatment had 8 replicates (n = 48). The nutrient treatments were applied following Güsewell (2005). N was supplied as KNO₃ and Ca(NO₃)₂, P was provided as KH₂PO₄. KNO₃ and KH₂PO₄ supplied part of the potassium, the rest was added as KCl. All the other essential macro- and micro-nutrients were provided in non-limiting supply using a standard Hoagland solution.

2.2.2 Reflectance measurements

The canopy reflectance of the grasses was measured inside the greenhouse under controlled conditions when plants were fully grown on February 19th 2016 using a FieldSpec Pro Fr spectroradiometer (Analytical Spectral Device, Boulder, GO, USA). This device measures reflectance between 350 and 2500 nm with a resampled spectral resolution of 1 nm. We performed the spectral measurements with several lamps positioned above the plants on each side of the spectroradiometer to ensure constant light conditions. The spectroradiometer was held by a tripod at nadir position approximately 20 cm above the canopy. The field of view was 8°, ensuring a ground field of view of approximately 12 cm². Each measurement of the spectroradiometer was the average of 50 successive scans. We measured each pot four times, after turning the pot by 90 degrees between each measurement to reduce geometrical effects. We averaged these four measurements to produce one canopy measurement per pot. The spectroradiometer was internally calibrated using a Spectralon white reference panel (Labsphere, North Sutton, NH, USA).

2.2.3 Leaf chemical measurements

Two leaves were sampled from each pot ($n = 48$) on February 25th 2016, collected in paper bags and dried in an oven at 60°C for 48h. The leaf samples were then ground using mixer mill (MM400, Retsch). The N concentration ($\text{g N } 100 \text{ g}^{-1}$ dry matter, %N) was measured using an elemental CN analyzer (Fisons NA 1500 NCS). The P concentration ($\text{g P } 100 \text{ g}^{-1}$ dry matter, %P) was measured using the total reflection X-ray fluorescence spectroscopy method (TXRF, S2 Picofox, Bruker, Germany). N:P ratios were calculated as the ratio between the weight based concentrations of N and P.

2.2.4 Spectral bands considered: original and resampled to satellite bands

Prior to analysis, the spectral range was reduced from 350-2500 nm to 400-2450 nm because the signal to noise ratio was too low for the minimum and maximum spectral ranges. This spectral range matches with the spectral coverage of earth observation sensors. We considered two sets of spectral bands: the original narrow band reflectance spectra and the spectra resampled to the properties of satellite sensor bands. We included the spectra resampled to satellite sensor bands to assess the possibility of detecting canopy N:P with VIs derived from actual satellite measurements. We included six different satellite sensors (MSI-Sentinel-2, OLCI-Sentinel-3, MODIS-Terra/Aqua, OLI-Landsat 8, WorldView 4, and RapidEye). We chose these sensors because they present a variety of band properties, from 5 to 21 bands and from 2.5 nm to 350 nm spectral resolution (Table 2.1, Figure 2.1). They are all operational and have a long measurement record in the case of MODIS and Landsat. We obtained the

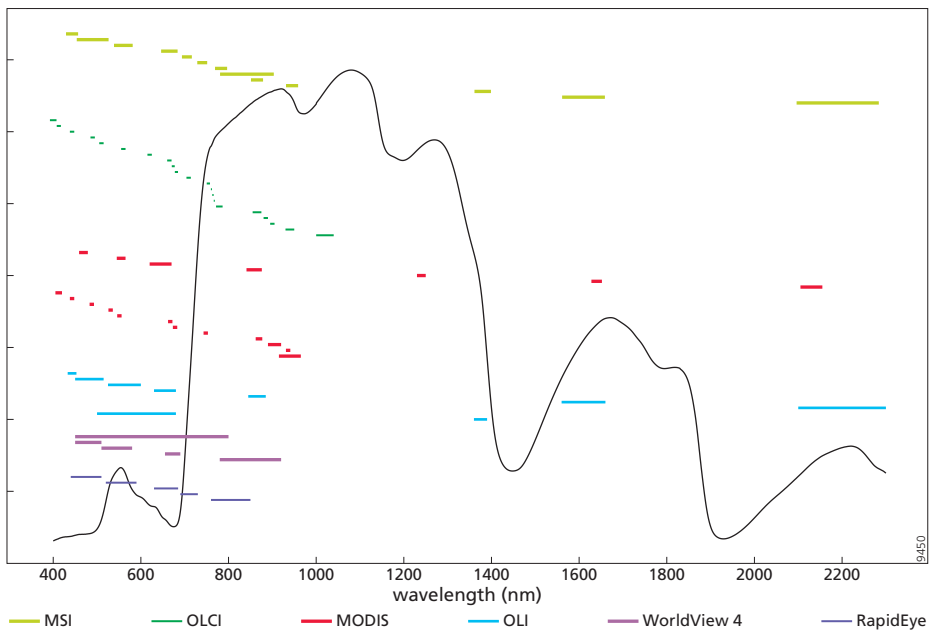


Figure 2.1 Spectral band position and bandwidth of the six satellite sensors included in the analysis (MSI, OLCI, MODIS, WorldView 4; RapidEye) projected on a Yorkshire fog (*Holcus lanatus*) reflectance spectrum measured using a spectroradiometer.

resampled spectra by using the spectral response function corresponding to each sensor band. The spectral response functions were downloaded from the website of the sensor producer. For the MODIS sensor, the measured spectral response function could not be found and was approximated using a standard normal distribution ($\mu = 0, \sigma = 1$).

Table 2.1 Properties of the six satellite sensors included in this study.

Satellite	Sensor	Number of bands	Bandwidth (nm)	Spatial resolution (m)
Sentinel 2	MSI	13	15-180	10-20-60
Sentinel 3	OLCI	21	2.5-40	300-1200
Terra-Aqua	MODIS	20	10-50	250-500-1000
Landsat 8	OLI	9	20-200	30
WorldView 4	WorldView 4	5	35-350	0.31-1.24
RapidEye	RapidEye	5	40-70	5

2.2.5 Data analysis

2.2.5.1 Descriptive statistics and vegetation indices

Descriptive statistics and boxplots of canopy N:P, N and P, hereafter designated as canopy traits, were produced and the Pearson correlation coefficients between each of these canopy traits were calculated.

2.2.5.2 Existing VIs

We evaluated existing VIs, which we selected through a literature search. We found 60 VIs to evaluate (Table A. 2.1), including VIs that were either developed for N or P detection or, when developed for another purpose, for instance VIs correlated with chlorophyll, photosynthesis or the presence of vegetation (see the ‘developed for’ column of the Table A. 2.1 for a detailed list), have been used for N or P detection.

Based on their equations, we categorized the VIs as either two bands simple ratio ($SR, \frac{band_1}{band_2}$), two bands normalized difference ($ND, \frac{band_1 - band_2}{band_1 + band_2}$) or three bands VI (TB). When they did not belong to any of the aforementioned categories, they were labelled other VI (OVI, see Table A. 2.1 for equations associated with TB or OVI).

We calculated linear models between canopy N:P, canopy N or canopy P and both the VIs and the natural log transformed VIs calculated using the original narrow band spectra.

The ten narrow band VIs that obtained the best correlation (r^2) with canopy N:P were also calculated using the spectra resampled to the six satellite sensor bands mentioned in section 2.4. The indices were computed using the sensor bands closest to the VI nominal wavelengths. Several VIs could not be computed because the sensors’ band locations were too distant (> 100 nm) from the VI nominal wavelength.

2.2.5.3 Optimized VIs

From the spectral signature we took from our experimental data, we calculated optimal VIs using combinations of all available wavelengths, i.e. a band combination analysis. Following Tian et al. (2011), every combination of two, λ_1 and λ_2 , or three wavelengths, λ_1 , λ_2 and λ_3 , was used to compute two bands or three bands VIs. The two bands VIs were categorized as simple

ratio (SR) $\frac{\lambda_1}{\lambda_2}$, simple difference (SD) $\lambda_1 - \lambda_2$ and normalized difference (ND) $\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}$, while the three bands indices were named TB1 $\frac{\lambda_1}{\lambda_2 + \lambda_3}$, TB2 $\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_3}$ and TB3 $\frac{\lambda_1 - \lambda_2 + 2 * \lambda_3}{\lambda_1 + \lambda_2 - 2 * \lambda_3}$. In order to decrease computation time, the three bands VIs were computed using only 1 out of 10 narrow bands (Pacheco-Labrador et al. 2014). Subsequently, linear regressions between each of the obtained VIs and canopy N:P, canopy N and canopy P were calculated. Specific wavelength regions of high correlation between the canopy traits and the three categories of two bands VI were identified using heatmaps. The optimized VIs were obtained by selecting the band combination with the highest r^2 , for both the narrow band spectra and the spectra resampled to satellite sensor bands.

2.2.5.4 Regression models

All regression models were assessed using the determination coefficient (r^2) and the cross-validated Relative Root Mean Squared Error (RRMSE_{cv}) values, obtained using the leave-one-out cross-validation method (Clevers and Gitelson 2013) and calculated following Equation 2.1 (Yao et al. 2010):

$$RRMSE_{cv} = \sqrt{\frac{1}{n} \times \sum_{i=1}^n (P_i - O_i)^2} \times \frac{1}{\bar{O}_i}, \quad (\text{Equation 2.1})$$

where $i = 1, 2, \dots, n$ is a measurement, with n the total number of measurements, P_i represents the predicted value, O_i , the observed value, and \bar{O}_i the mean of all observed values. The significance level was at a p -value < 0.01 . All the statistical analyses were performed in the R environment (R Development Core Team 2014).

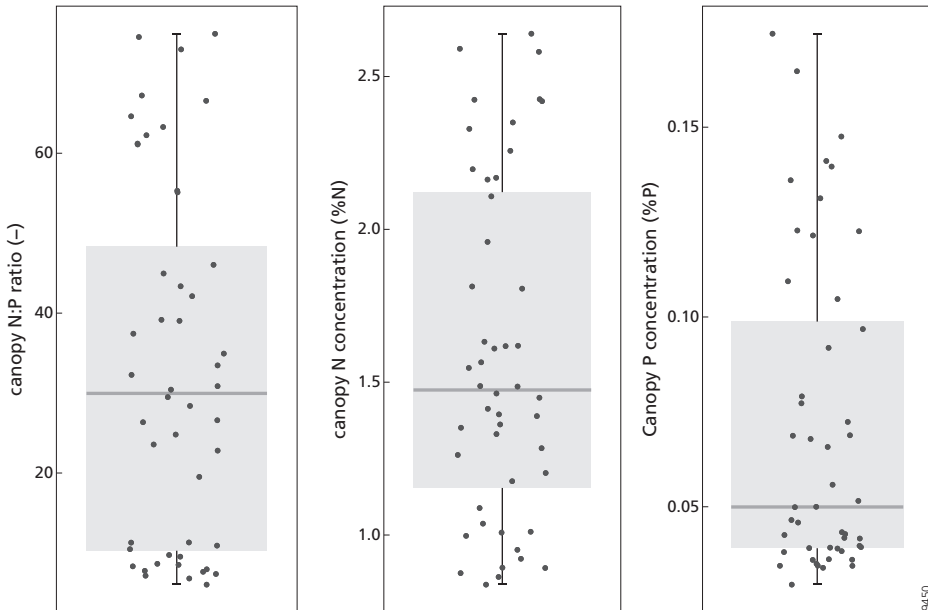


Figure 2.2. Boxplot of canopy N:P (-), canopy N (%N) and canopy P (%P) in *Holcus lanatus* L. grasses ($n = 48$).

2.3 Results

2.3.1 Descriptive statistics of canopy N:P, canopy N and canopy P

The results of the descriptive statistics showed that the range of canopy N:P values (6.1-75.0, Figure 2.2) corresponds to the range usually observed in natural environments (Roeling et al. 2018). As expected, canopy N:P was correlated with both canopy N and canopy P (Table 2.2)

Table 2.2. Pearson correlation matrix between canopy N, canopy P and canopy N:P. The values in the lower part of the table correspond to the correlation coefficients while the upper part of the table corresponds to the p-values.

	canopy N	canopy P	canopy N:P
canopy N	/	<i>p</i> -value = 0.000	<i>p</i> -value = 0.000
canopy P	<i>r</i> = -0.66	/	<i>p</i> -value = 0.000
canopy N:P	<i>r</i> = 0.92	<i>r</i> = -0.81	/

2.3.2 Original narrow band spectra

2.3.2.1 Existing vegetation indices

Among the 60 existing VIs tested, 31 VIs showed significant relationships with at least one canopy trait, while 29 VIs showed no significant relationship with any of the canopy traits (Table A. 2.1). The VI showing the highest correlation with canopy N:P ($r^2 = 0.48$), was the TB (R_{498} , R_{413} , R_{442} , Figure 2.3) VI developed for N detection (Tian et al. 2011). The results obtained for canopy N and canopy P were similar, with the highest r^2 equal to 0.44 for canopy N with ND (R_{1220} , R_{710}) VI. For canopy P, the highest correlation ($r^2 = 0.52$) was obtained with TB (R_{498} , R_{413} , R_{442}), hence the same VI as for canopy N:P. The two VIs showing the highest correlation with canopy N:P, TB (R_{498} , R_{413} , R_{442}) and TB (R_{434} , R_{496} , R_{401}) were both based

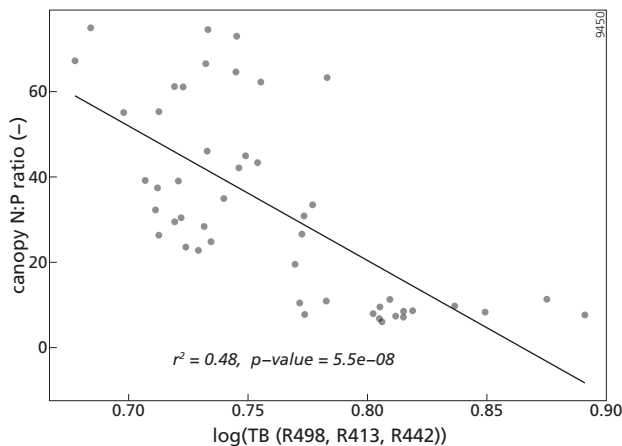


Figure 2.3. Scatterplot and regression line between canopy N:P and the TB (R_{498} , R_{413} , R_{442}) vegetation index, i.e. the existing vegetation index with the highest correlation with canopy N:P ($r^2 = 0.48$, $n = 48$). The vegetation index was calculated using the original narrow band spectra.

on three wavelengths located in the blue region of the spectrum. Other VIs that showed a significant relationship with canopy N:P were based on the NIR and red-edge regions of the spectrum.

2.3.2.2 Optimized vegetation indices

We used heatmaps to investigate the wavelength regions of high correlation between the two bands VIs, obtained from the band combination analysis, and the three canopy traits (Figure 2.4). A first region of high correlation with canopy N:P (r^2 between 0.4 and 0.7) was located

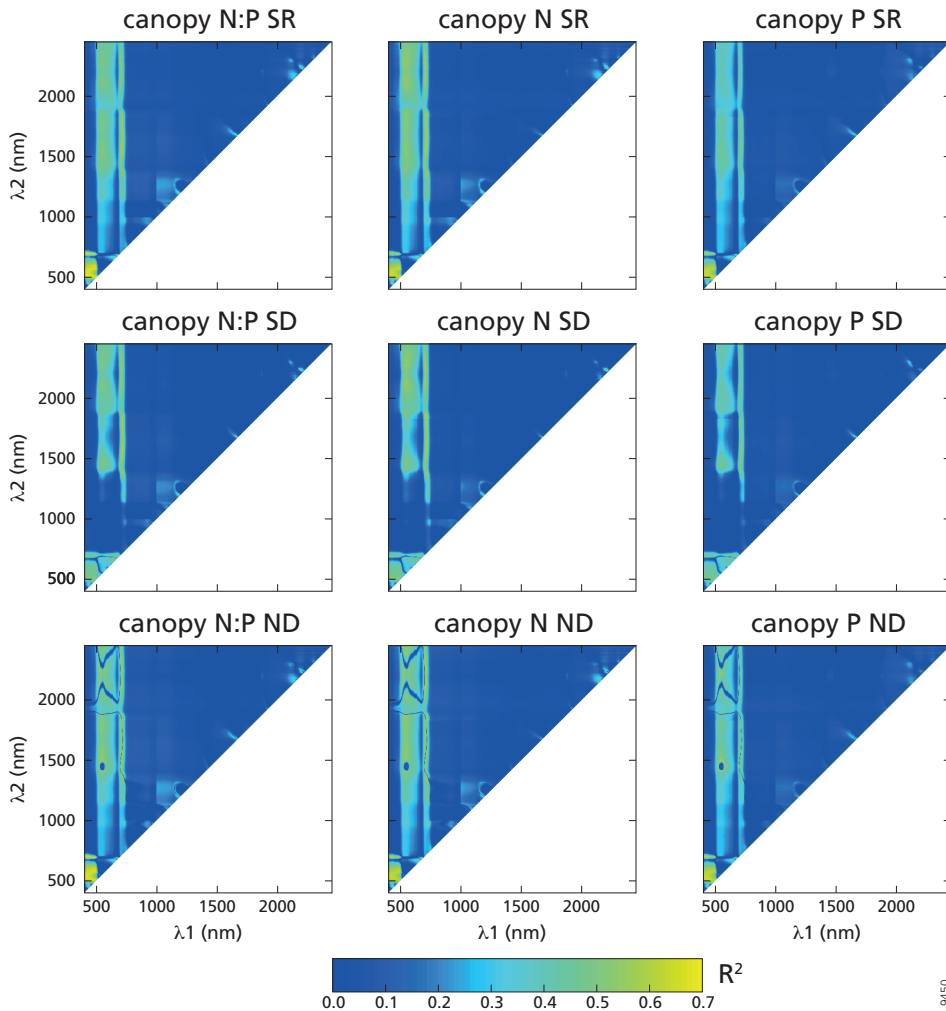


Figure 2.4 Heatmaps showing the coefficient of determination (r^2) between the canopy trait and each combination of two bands (λ_1 and λ_2 , nm) between 400-2450 nm for canopy N:P, canopy N and canopy P and each VI category investigated. SR, Simple Ratio; SD, Simple Difference, ND, Normalized Difference.

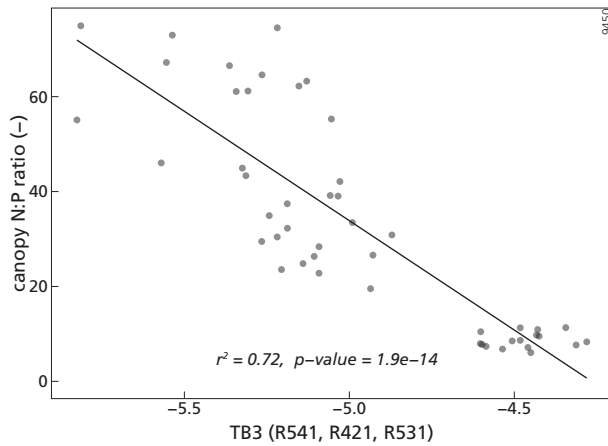


Figure 2.5 Scatterplot and regression line between canopy N:P and the TB3 (R_{541} , R_{421} , R_{531}) vegetation index, i.e. the optimized vegetation index with the highest correlation with canopy N:P ($r^2 = 0.72$, $n = 48$). The vegetation index was calculated using the original narrow band spectra.

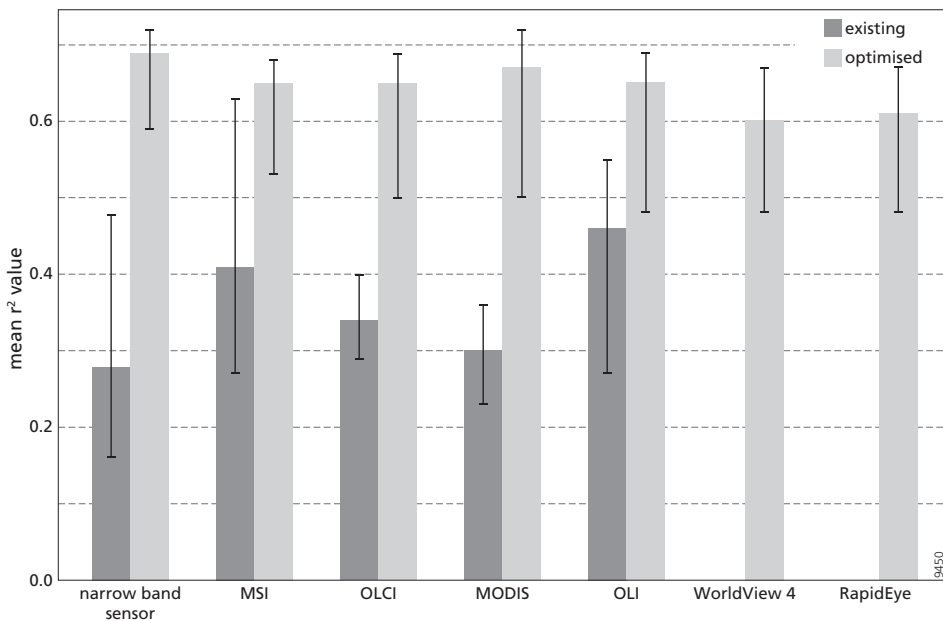


Figure 2.6 Mean r^2 value of the relationship between canopy N:P and the vegetation indices (VIs) for the seven sensors included in this analysis and for both the existing and optimized VIs. The vertical lines represent the minimum and maximum r^2 obtained for each group. No existing VI was calculated for the WorldView 4 and RapidEye sensors because their band positions are too distant (> 100 nm) from the nominal VIs' wavelengths.

in the blue region for (400-500 nm) and in the green-red region for (500-700 nm). A second region (r^2 between 0.4 and 0.6) was located in the green-red region for (500-700 nm) and in the SWIR region for (1400-2450 nm). The regions of high correlation with canopy N:P were similar across the VIs categories investigated, i.e. SR, SD and ND, and also had high correlations for canopy N and canopy P.

The optimized VI TB3 (R_{541} , R_{421} , R_{531} , Figure 2.5), located in the blue and green region of the spectrum showed the highest correlation with canopy N:P and had a higher r^2 ($r^2 = 0.72$, Table A. 2.2) than the best performing optimized VIs for canopy N ($r^2 = 0.69$; TB3, R_{431} , R_{2431} , R_{561}) and for canopy P ($r^2 = 0.67$; TB1, R_{531} , R_{541} , R_{421}).

Regarding the influence of the VI category on the result, the optimized VIs from the SD category, which were composed of a combination of bands from the red-edge and SWIR regions, showed lower correlations ($r^2 = 0.53$ - 0.59) than all the other VI categories ($r^2 = 0.64$ - 0.72), mostly composed of a combination of blue and green bands. This result holds for all canopy traits (canopy N:P, N and P).

Compared to the results obtained for the existing VIs, the optimized VIs performed better regarding r^2 values (Figure 2.6) for canopy N:P. The best optimized VI explained 24 percent more of the variation in canopy N:P values compared to the best performing existing VI.

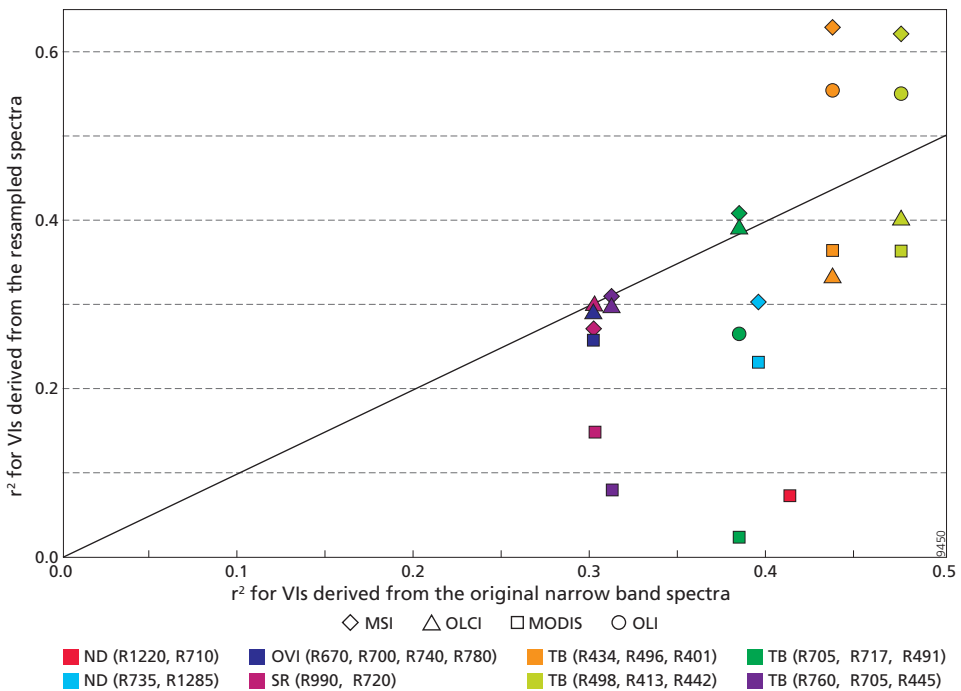


Figure 2.7 Comparison between the r^2 obtained for the relationships between canopy N:P and the existing VIs calculated using either the original narrow band spectra or the spectra resample to satellite sensors' band properties (MSI, OLCI, MODIS, OLI). No existing VI was calculated for the RapidEye and WorldView 4 sensors because their band positions are too distant (> 100 nm) from the nominal VIs' wavelengths.

2.3.3 Spectra resampled to satellite sensors' bands

2.3.3.1 Existing vegetation indices

The TB (R_{434} , R_{496} , R_{401}) VI calculated for the MSI sensor aboard Sentinel 2 showed the highest correlation with canopy N:P ($r^2 = 0.63$), canopy N ($r^2 = 0.58$) and canopy P ($r^2 = 0.62$) compared to the other existing VIs tested (Table A. 2.3). Regarding the influence of the sensor band properties on the result, the existing VIs calculated using the MSI and OLI sensors band showed in general higher correlation with canopy N:P than the same VIs calculated based on OLCI or MODIS sensors (Figure 2.7). For the existing VIs TB (R_{498} , R_{413} , R_{442}), TB (R_{434} , R_{496} , R_{401}) and TB (R_{705} , R_{717} , R_{491}) calculated based on MSI sensor bands, the TB (R_{498} , R_{413} , R_{442}) and TB (R_{434} , R_{496} , R_{401}) VIs calculated based on OLI sensor bands as well as for the TB (R_{705} , R_{717} , R_{491}) VI calculated based on OLCI sensor properties, the obtained r^2 were higher than for the same VIs calculated with the original narrow band spectra. Similar results were observed for canopy N and canopy P. On the contrary, all the existing VIs calculated with the spectra resampled to MODIS bands showed lower r^2 with canopy N:P compared to their narrow band spectra counterpart and three existing VIs showed a non-significant relationship with canopy N:P.

2.3.3.1 Optimized vegetation indices

Canopy N:P could be related to the optimized VIs calculated from the resampled spectra with r^2 values ranging from 0.48 to 0.72 (Table A. 2.2). Similar results were obtained for canopy N and canopy P. The SR, ND and TB categories of optimized VIs mostly performed better with canopy N:P (r^2 between 0.55-0.72) compared to the SD category of VI (r^2 between 0.48-0.59). Among the bands selected for the SR, ND and TB categories of optimized VIs, the blue and green regions of the spectrum were dominant, similar to what was observed for narrow bands optimized VIs.

The difference in satellite sensors band properties did not greatly affect the results as the r^2 values were in the same range for all sensors per VI category. The satellite sensors WorldView 4 and RapidEye, both with five bands and bandwidths between 35-140 nm, showed r^2 between 0.55 and 0.67 with canopy N:P for the SR, ND and TB categories of VI. This is comparable to the results obtained for satellites sensors with a higher number of bands and narrower bands, e.g. MSI and MODIS, for which the obtained r^2 values were between 0.67 and 0.72 for the same VI categories.

When comparing the performance of existing and optimized VIs for canopy N:P detection (Figure 2.6), optimized VIs generally performed better than existing VIs for all satellite sensors tested. However, the existing VIs calculated for the MSI and OLI sensors showed r^2 values in an overlapping range with their optimized counterparts.

2.4 Discussion

2.4.1 Original narrow band spectra

2.4.1.1 Existing vegetation indices

Existing VIs performed similarly for canopy N:P detection as they did for canopy N and canopy P regarding obtained r^2 values. More specifically, the two best performing VIs for

canopy N:P in this study, TB (R_{498} , R_{413} , R_{442}) and TB (R_{434} , R_{496} , R_{401}), both based on the blue region of the spectrum, proved to be interesting candidates for canopy N:P detection. In previous studies, these VIs were highly correlated to canopy N in rice crops ($r^2 = 0.81$ and $r^2 = 0.84$, respectively (Tian et al. (2011))), but showed a weak correlation with canopy N in Holm oak leaves (Pacheco-Labrador et al. 2014).

Among the 60 VIs tested, almost 50% ($n = 29$) failed to show a significant relationship with canopy N:P, canopy N or canopy P while previous studies found a significant relationship. This unreliability in prediction accuracy for previously validated VIs was already observed by Pacheco-Labrador et al. (2014), who found that the majority of the published vegetation indices tested could not be correlated to canopy N in Holm oak leaves. This might be explained by differences in growth conditions and species investigated as well as by the influence of the range in canopy N or canopy P values. This highlights the influence of the dataset on the prediction accuracy.

2.4.1.2 Optimized vegetation indices

The optimized VIs developed specifically for this study performed better for canopy N:P detection compared to the existing VIs tested, for both r^2 and $RRMSE_{cv}$ values.

The SR, ND and TB categories of optimized VIs were based on a combination of wavelengths from the blue and green regions of the spectrum. The blue region is linked to pigment absorption, which peaks at 430 nm for chlorophyll a (chlor-a) and 460 nm for chlorophyll b (chlor-b) (Kumar et al. 2006). The SD, ND and TB₃ optimized VIs for canopy N:P included wavelengths located at 427 nm and 421 nm, hence close to one of the absorption peaks of chlor-a. Other indices located in the blue and green regions of the spectrum have been related to chlor-a and chlor-b concentration in *Vitis vinifera* leaves (the BGI2 index, SR, R_{450} , R_{550} ; Zarco-Tejada et al. (2005)) or to canopy N in rice (ND, R_{573} , R_{444} ; Tian et al. (2011)). The combination of wavelengths from the blue and green regions might thus be linked with chlorophyll and has also been explored for canopy N detection (Tian et al. 2011). Our results showed that this wavelength combination is correlated with canopy N:P, canopy N and canopy P.

The optimized SD VI for canopy N:P combined wavelengths located in the red-edge (718 nm) and in the SWIR regions (1577 nm). The SWIR region is characterized by several N and protein absorption features. The absorption features located at 1500 nm and 1510 nm might explain the second wavelength (1577 nm) being selected, because the absorption features are known to be broadened due to scattering (Kumar et al. 2006). However, the SWIR region of the spectra is influenced by the absorption features of many compounds that interfere with each other (Kumar et al. 2006) which render the signal difficult to interpret. In particular, the strong absorption due to water molecules greatly influences the reflectance. This might be the reason why the proportion of variance explained by the SD category of optimized VIs ($r^2 = 0.53-0.59$) was lower than for the SR, ND and TB categories of optimized VIs ($r^2 = 0.64-0.72$).

2.4.2 Spectra resampled to satellite sensors' bands

2.4.2.1 Existing vegetation indices

The correlation between canopy N:P and the existing VIs was influenced by the sensors' band properties because different sensors showed different correlation with canopy N:P for the same

VIs. Two VIs (TB R_{498} , R_{413} , R_{442} and TB R_{434} , R_{496} , R_{401}) calculated from MSI and OLI sensors showed higher correlation ($r^2 = 0.55-0.63$) than when calculated from the OLCI and MODIS sensors ($r^2 = 0.36-0.40$) or even narrow band sensors ($r^2 = 0.44-0.48$) (Figure 2.7). The bands from MSI and OLI sensors used for the VIs calculation were broader than their counterparts from the OLCI and MODIS sensors and hence from the original narrow band spectra. The bandwidth of the sensors thus influenced the correlation obtained for some of the existing VIs.

2.4.2.2 Optimized vegetation indices

The correlation between canopy N:P and the optimized VIs was not greatly influenced by the sensors' band properties and the correlation was stable across the different satellite sensors tested. Although this is in contradiction to what was observed for existing VIs, this might indicate that broad band sensors, such as RapidEye and WorldView 4, could be useful for canopy N:P detection. However, our results consistently showed that the blue region was important for canopy N:P detection. Although VIs based on the blue region of the spectrum do not represent a difficulty for in situ studies, including the blue region for VIs calculated from satellite sensors might lead to an interpretation problem due to atmosphere Rayleigh scattering.

2.4.3 Future perspectives

This analysis investigated the possibility of detecting canopy N:P using VIs compared to canopy N and canopy P detection as well as the influence of the sensors band properties on the relationships. The results obtained in this analysis might have concrete in-situ application perspectives in the field of ecology given the importance of the canopy N:P for biodiversity studies (Fujita et al. 2014; Roeling et al. 2018; Wassen et al. 2005). Spectral VIs might be a useful method to detect canopy N:P in grasslands in a non-destructive and time efficient manner that would, for example, allow to monitor the seasonal evolution of canopy N:P or trends in changing N:P ratios in response to eutrophication or other global change factors (Wassen et al. 2013). However, as with remote sensing of canopy N or canopy P, remote sensing of canopy N:P is not a direct measurement of canopy N:P values. Especially, the correlation between canopy N:P and both canopy N and canopy P (Table 2.2) renders it difficult to distinguish between the separate influences on the reflectance signal. Moreover, similar studies investigating canopy N:P detection should be done on more plant species and plant communities in order to validate the results obtained in this study. Further studies should also investigate the influence of the spatial resolution of the satellite sensors on canopy N:P detection as this needs to be studied before actual satellite sensors measurements can be exploited to sense canopy N:P in natural environment.

2.5 Conclusion

Canopy N remote detection has already been extensively studied using both spectroradiometer and satellite measurements. On the contrary, canopy N:P, despite being an important indicator of nutrient limitation, has seldom been studied with remote sensing techniques. In this study, we investigated the possibility of detecting canopy N:P in the common grass species *Holcus*

lanatus using VIs developed for canopy N and canopy P detection. The results showed that using VIs for canopy N:P detection was as effective as when applied for canopy N or canopy P detection. This held for both existing and optimized VIs as well as for narrow band and broader band VIs calculated from the spectra resampled to the spectral properties of six different satellite sensors. The influence of different satellite sensors' band properties was unclear as it differed between existing and optimized VIs. Existing VIs calculated with MSI and OLI sensors bands showed higher correlation with canopy N:P compared to the other sensors tested. In contrast, the correlation with optimized VIs was not affected by the differences in sensors' bands. Satellite sensors with a limited number of broad bands, such as WorldView 4 and RapidEye, yielded similar results as sensors with multiple and narrower bands, like MSI or Sentinel 3. In the future, these results might lead to practical applications using handheld spectrometers for in situ canopy N:P detection in grasslands. The observed consistent importance of the blue region of the spectrum for canopy N:P detection might render canopy N:P detection with actual satellite sensors complicated due to the interference with the Rayleigh scattering.

2.6 Acknowledgments

This work was supported by The Netherlands Organization for Scientific Research (NWO) [NWO ALW-GO-AO/14-12]. We would like to acknowledge Ton Markus for his help with the figures presented in this paper.

2.7 Appendix

Table A.2.1. Relationship between the existing vegetation indices (VIs) ($n = 60$) and canopy N:P, canopy N and canopy P. Coefficient of determination (r^2), p-value and Relative Root Mean Squared Error of cross-validation (RRMSEcv). For each VI, the best model, linear (L) or logarithmic (LOG), based on r^2 value, is presented. The VIs were classified into four categories: ND = normalized difference; SR = simple ratio; TB = three band; OVI = other vegetation index. The VIs are sorted from the highest to the lowest correlation with canopy N:P.

VI	Canopy N:P			Canopy N			Canopy P			Equation	Developed for	References			
	Model	r^2	p-value	RRMSEcv	Model	r^2	p-value	RRMSEcv	Model				r^2	p-value	RRMSEcv
TBVI (R498, R413, R442)	LOG	0.48	0.000	0.50	LOG	0.42	0.000	0.27	LOG	0.52	0.000	0.43	(R498 + R413)/R442	N detection	(Tian et al. 2011)
TBVI (R434, R496, R401)	L	0.44	0.000	0.52	L	0.38	0.000	0.28	LOG	0.48	0.000	0.44	R434/(R496 + R401)	N detection	(Tian et al. 2011)
NDVI (R1220, R710)	L	0.41	0.000	0.54	LOG	0.44	0.000	0.26	LOG	0.37	0.000	0.48	(R1220 - R710)/(R1220 + R710)	N detection	(Zhu et al. 2007a)
NDVI (R693, R1770)	L	0.40	0.000	0.54	L	0.39	0.000	0.28	L	0.34	0.000	0.49	(R693-R1770)/(R693 + R1770)	N detection	(Ferwerda et al. 2005)
NDVI (R735, R1285)	L	0.40	0.000	0.54	L	0.44	0.000	0.27	L	0.30	0.000	0.50	(R735 - R1285)/(R735 + R1285)	N detection	(Pacheco-Labrador et al. 2014)
TBVI (R705, R717, R491)	L	0.38	0.000	0.55	L	0.37	0.000	0.28	L	0.36	0.000	0.48	R705/(R717 + R491)	N detection	(Tian et al. 2011)
TBVI (R760, R705, R445)	L	0.31	0.000	0.58	LOG	0.32	0.000	0.29	LOG	0.32	0.000	0.49	(R760-R705)/(R760 + R705-2*R445)	N detection	(Tian et al. 2011)
TBVI (R760, R723, R2387)	LOG	0.31	0.000	0.58	LOG	0.34	0.000	0.29	LOG	0.30	0.000	0.50	(R760-R2387)/(R723-R2387)	P detection	(Wang et al. 2016a)
SR (R990, R720)	LOG	0.30	0.000	0.59	LOG	0.35	0.000	0.28	LOG	0.28	0.000	0.51	R 990/R720	N detection	(Yao et al. 2010)
OVI (R670, R700, R740, R780) REP	LOG	0.30	0.000	0.58	LOG	0.33	0.000	0.29	LOG	0.33	0.000	0.49	700 + 40 * ((R670 + R780)/(2-R700))/(R740 - R700)	red-edge	(Horler et al. 1983)
TBVI (R924, R703, R423)	LOG	0.30	0.000	0.58	LOG	0.32	0.000	0.29	LOG	0.30	0.000	0.50	(R924-R703 + 2*R423)/(R924 + R703 + 2*R423)	N detection	(Wang et al. 2012)

VI	Canopy N:P		Canopy N		Canopy P		Equation	Developed for	References						
	Model	r ²	p-value	RRMSEcv	Model	r ²				p-value	RRMSEcv				
SR (R1220, R610)	LOG	0.30	0.000	0.58	LOG	0.31	0.000	0.29	LOG	0.26	0.000	0.52	R1220/R610	N detection	(Zhu et al. 2007b)
TBVI (R754, R709, R681) MTCI	LOG	0.29	0.000	0.59	LOG	0.31	0.000	0.29	LOG	0.31	0.000	0.50	(R754 – R709)/(R709 – R681)	red-edge-chlorophyll	(Dash and Curran 2004)
SR (R550, R800)	L	0.27	0.000	0.59	L	0.30	0.000	0.29	L	0.29	0.000	0.51	R550/R 800	chlorophyll	(Buschmann and Nagel 1993)
TBVI (R1310, R1720, R730)	LOG	0.27	0.000	0.60	LOG	0.33	0.000	0.29	LOG	0.22	0.001	0.53	R1310/(R1720 + R730)	N detection	(Pacheco-Labrador et al. 2014)
NDVI (R850, R550) Green NDVI	LOG	0.27	0.000	0.59	LOG	0.30	0.000	0.29	LOG	0.28	0.000	0.51	(R850-R550)/(R850 + R550)	chlorophyll	(Gitelson et al. 1996)
OVI (R790, R550) Clgreen	LOG	0.27	0.000	0.60	LOG	0.30	0.000	0.30	LOG	0.28	0.000	0.51	R790/R550-1	red-edge-chlorophyll	(Gitelson et al. 2005)
TBVI (R720, R860, R450)	L	0.26	0.000	0.60	L	0.29	0.000	0.30	L	0.27	0.000	0.51	(R720-R450)/(R860-R450)	chlorophyll	(le Maire et al. 2008)
SR (R810, R560)	LOG	0.25	0.000	0.60	LOG	0.28	0.000	0.30	LOG	0.26	0.000	0.52	R 810/R560	N detection	(Xue et al. 2004)
NDVI (R750, R705)	L	0.24	0.000	0.61	LOG	0.25	0.000	0.31	LOG	0.26	0.000	0.52	(R750-705)/(R750 + 705)	N detection	(Corp et al. 2010)
NDVI (R790, R720) NDRE	LOG	0.24	0.000	0.61	LOG	0.27	0.000	0.30	LOG	0.27	0.000	0.51	(R790 – R720)/(R790 + R720)	red-edge	(Fitzgerald et al. 2010)
OVI (R790, R720) Clred-edge	LOG	0.24	0.000	0.61	LOG	0.26	0.000	0.30	LOG	0.26	0.000	0.52	(R790/R720)-1	red-edge-chlorophyll	(Gitelson et al. 2005)
NDVI (R860, R720)	LOG	0.23	0.001	0.61	LOG	0.27	0.000	0.30	LOG	0.25	0.000	0.52	(R860-R720)/(R860 + R720)	N detection	(Yao et al. 2010)
TBVI (R760, R734, R834)	LOG	0.22	0.001	0.62	LOG	0.22	0.001	0.31	LOG	0.20	0.001	0.54	R760/(R734 + R834)	N detection	(Tian et al. 2011)
TBVI (R715, R860, R445)	L	0.20	0.001	0.63	L	0.23	0.001	0.31	L	0.21	0.001	0.54	(R715-860)/(R715 + R860 + 2*R445)	chlorophyll	(le Maire et al. 2008)

VI	Canopy N:P		Canopy N		Canopy P		Equation	Developed for	References						
	Model	r ²	p-value	RRMSEcv	Model	r ²				p-value	RRMSEcv				
TBVI (R750, R705, R445)	L	0.20	0.002	0.63	L	0.21	0.001	0.32	LOG	0.21	0.001	0.53	(R750-705)/(R750 + R705 + 2*R445)	chlorophyll	(Sims and Gamon 2002)
TBVI (R708, R710, R680)	L	0.17	0.003	0.64	L	0.16	0.005	0.33	L	0.15	0.007	0.55	(R708-R710)/(R708-R680)	N detection	(Tian et al. 2011)
TBVI (R730, R720, R725)	L	0.17	0.004	0.64	L	0.21	0.001	0.32	L	0.20	0.002	0.54	(R730-R725)/(R720-R725)	P detection	(Wang et al. 2016a)
TBVI (R720, R729, R726)	LOG	0.17	0.004	0.64	LOG	0.20	0.001	0.32	LOG	0.19	0.002	0.54	(R720-R729)/(R720 + R729 - 2*R726)	P detection	(Wang et al. 2016a)
SR (R800, R635) PSSRb	L	0.16	0.005	0.64	LOG	0.17	0.004	0.33	LOG	0.15	0.006	0.55	R 800/R 635	photosynthesis	(Blackburn 1998)
NDVI (R544, R551)	L	0.16	0.005	0.64	L	0.17	0.004	0.33	L	0.16	0.005	0.55	(R544-R551)/(R544 + R551)	P detection	(Kawamura et al. 2011)
OVI (R830, R660, R670) RDVI	L	0.12	0.018	0.66	L	0.12	0.017	0.34	L	0.09	0.035	0.57	((R830/R660) - 1)/sqrt((R830/R670) + 1)	vegetation presence	(Chen 1996)
NDVI (R1080, R1460)	LOG	0.11	0.020	0.66	LOG	0.09	0.043	0.34	LOG	0.10	0.032	0.57	(R1080-R1460)/(R1080 + 1460)	P detection	(Mahajan et al. 2014)
SR (R950, R660)	L	0.11	0.022	0.66	L	0.11	0.020	0.34	LOG	0.08	0.057	0.58	R950/R 660	N detection	(Zhu et al. 2007a)
NDVI (R870, R1450)	LOG	0.10	0.025	0.66	LOG	0.09	0.044	0.34	LOG	0.07	0.063	0.58	(R870-R1450)/(R870 + R1450)	P detection	(Pimstein et al. 2011)
SR (R830, R660) RVI	L	0.09	0.034	0.67	L	0.09	0.034	0.34	L	0.07	0.075	0.58	R830/R660	red-edge	(Jordan 1969)
NDVI (R1050, R1100)	L	0.09	0.035	0.67	L	0.18	0.003	0.32	L	0.05	0.136	0.59	(R1050-R1100)/(R1050 + R1100)	N detection	(Pacheco-Labrador et al. 2014)
SD (R533, R565)	L	0.08	0.045	0.67	L	0.08	0.055	0.34	L	0.07	0.066	0.58	R533-R565	N detection	(Tian et al. 2011)
NDVI (R1645, R1715)	L	0.08	0.048	0.68	L	0.07	0.068	0.34	L	0.09	0.035	0.58	(R1645-R1715)/(R1645 + R 1715)	P detection	(Pimstein et al. 2011)

VI	Canopy N:P		Canopy N		Canopy P		Equation	Developed for	References						
	Model	r ²	p-value	RRMSEcv	Model	r ²				p-value	RRMSEcv				
NDVI (R1680, R1510) NDNI	LOG	0.07	0.075	0.68	LOG	0.04	0.202	0.35	LOG	0.07	0.064	0.58	(R1680-R1510)/(R1680 + R1510)	N detection	(Serrano et al. 2002)
SR (R950, R680)	L	0.07	0.079	0.68	L	0.07	0.079	0.34	L	0.04	0.166	0.59	R950/R680	N detection	(Zhu et al. 2007a)
NDVI (R830, R660) NDVI	L	0.06	0.089	0.69	L	0.06	0.086	0.35	L	0.06	0.101	0.58	(R830-R660)/(R830 + R660)	vegetation presence	(Huete and Jackson 1987)
SR (R800, R680) PSSRa	L	0.05	0.111	0.69	L	0.05	0.121	0.35	L	0.04	0.202	0.59	R800/R680	photosynthesis	(Blackburn 1998)
NDVI (R915, R920)	L	0.05	0.116	0.69	L	0.07	0.076	0.35	L	0.04	0.201	0.59	(R915-R920)/(R915 + R920)	N detection	(Pacheco-Labrador et al. 2014)
NDVI (R825, R1120)	L	0.04	0.189	0.69	L	0.05	0.140	0.35	L	0.01	0.573	0.60	(R825-R1120)/(R825 + R1120)	N detection	(Pacheco-Labrador et al. 2014)
ND (R533, R565)	L	0.02	0.292	0.70	L	0.03	0.222	0.36	L	0.03	0.240	0.59	abs(R533 - R565)/(R533 + R565)	N detection	(Tian et al. 2011)
SR (R533, R565)	LOG	0.02	0.292	0.70	LOG	0.03	0.222	0.36	LOG	0.03	0.240	0.59	R533/R565	N detection	(Tian et al. 2011)
SR (R890, R437)	LOG	0.02	0.352	0.70	LOG	0.01	0.592	0.36	L	0.01	0.418	0.60	(R890/R437)	N detection	(Mirnik et al. 2005)
NDVI (R653, R688)	L	0.01	0.496	0.70	L	0.03	0.260	0.35	L	0.01	0.449	0.61	(R653-R688)/(R653 + R688)	P detection	(Kawamura et al. 2011)
NDVI (R531, R570) PRI	L	0.01	0.540	0.71	L	0.01	0.444	0.36	L	0.01	0.452	0.60	(R531-R570)/(R531 + R570)	photosynthesis	(Gamon et al. 1997)
TBVI (R800, R445, R680) SIPI	L	0.00	0.713	0.71	L	0.00	0.645	0.36	L	0.00	0.699	0.60	(R800 -R445)/(R800 -R680)	chlorophyll	(Peñuelas et al. 1995)
SR (R1129, R462)	LOG	0.00	0.727	0.70	LOG	0.01	0.470	0.36	LOG	0.00	0.792	0.60	(1129/R462)	P detection	(Mirnik et al. 2005)
TBVI (R830, R660, R670) SAVI	LOG	0.00	0.739	0.72	LOG	0.00	0.884	0.37	LOG	0.00	0.854	0.60	1.5*(R830-R660)/(R830 +R670 + 0.5)	vegetation presence	(Huete 1988)

VI	Canopy N:P			Canopy N			Canopy P			Equation	Developed for	References			
	Model	r ²	p-value	RRMSEcv	Model	r ²	p-value	RRMSEcv	r ²				p-value	RRMSEcv	
NDVI (R573, R540)	L	0.00	0.804	0.71	L	0.00	0.994	0.36	L	0.00	0.967	0.60	(R573-R540)/(R573+R540)	N detection	(Hansen and Schjoerring 2003)
OVI (R830, R660) RDVI	LOG	0.00	0.828	0.72	L	0.00	0.862	0.36	LOG	0.00	0.934	0.60	(R830-R660)/ sqrt(R830 + R660)	photosynthesis	(Roujean and Breon 1995)
NDVI (R662, R686)	L	0.00	0.829	0.70	L	0.00	0.846	0.36	L	0.00	0.645	0.61	(R662-R686)/(R662 + R686)	P detection	(Kawamura et al. 2011)
NDVI (R523, R583)	L	0.00	0.857	0.71	L	0.00	0.708	0.36	L	0.00	0.670	0.60	(R523-R583)/(R523 + R583)	P detection	(Kawamura et al. 2011)
OVI (R800, R670) OSAVI	L	0.00	0.890	0.72	L	0.00	0.818	0.37	L	0.00	0.872	0.60	(1 + 0.16) * (R800-R670)/(R800 + R670 + 0.16)	vegetation presence	(Rondeaux et al. 1996)
SR (R800, R470) PSSRC	L	0.00	0.909	0.71	L	0.00	0.673	0.36	L	0.00	0.878	0.60	R800/R470	photosynthesis	(Blackburn 1998)
NDVI (R1200, R1290)	L	0.00	0.999	0.70	L	0.01	0.544	0.35	L	0.00	0.706	0.60	(R1200-R1290)/ (R1200 + R1290)	N detection	(Pacheco-Labrador et al. 2014)

Table A. 2.2. Optimized vegetation indices for each canopy trait and each VI category (SR, SD, ND, TB1, TB2 and TB3). Bands composing the VI (λ_1 , λ_2 and λ_3 , nm), coefficient of determination (r^2) and the Relative Root Mean Squared error of cross-validation (RRMSEcv) are showed. The results are showed for the original narrow band spectra and the spectra resampled to satellite sensors bands (MSI, aboard Sentinel 2, OLCI, aboard Sentinel 3, MODIS aboard Terra-Aqua, and OLI, aboard Landsat 8). SD = Simple difference; SR = Simple Ratio; ND = Normalized Difference, TB = Three Bands indices, PAN = Panchromatic band.

sensor	canopy trait	VI category	λ_1 (nm)	λ_2 (nm)	λ_3 (nm)	r^2	p-value	RRMSEcv
narrow band sensor	Canopy N:P	SR	427	524	\	0.71	0.000	0.38
	Canopy N:P	SD	718	1577	\	0.59	0.000	0.45
	Canopy N:P	ND	427	524	\	0.70	0.000	0.39
	Canopy N:P	TB1	491	661	531	0.71	0.000	0.37
	Canopy N:P	TB2	531	411	541	0.70	0.000	0.38
	Canopy N:P	TB3	541	421	531	0.72	0.000	0.37
	Canopy N	SR	428	529	\	0.67	0.000	0.20
	Canopy N	SD	717	1833	\	0.58	0.000	0.23
	Canopy N	ND	428	529	\	0.66	0.000	0.21
	Canopy N	TB1	531	541	411	0.67	0.000	0.20
	Canopy N	TB2	531	411	551	0.67	0.000	0.20
	Canopy N	TB3	431	2431	561	0.69	0.000	0.20
	Canopy P	SR	422	524	\	0.66	0.000	0.36
	Canopy P	SD	717	1553	\	0.53	0.000	0.42
	Canopy P	ND	422	524	\	0.66	0.000	0.35
	Canopy P	TB1	531	541	421	0.67	0.000	0.35
Canopy P	TB2	421	531	541	0.66	0.000	0.35	
Canopy P	TB3	721	521	451	0.64	0.000	0.36	
MSI	Canopy N:P	SR	443	560	\	0.67	0.000	0.40
	Canopy N:P	SD	2190	705	\	0.53	0.000	0.48
	Canopy N:P	ND	443	560	\	0.67	0.000	0.40
	Canopy N:P	TB1	443	560	490	0.67	0.000	0.40
	Canopy N:P	TB2	560	443	490	0.66	0.000	0.41
	Canopy N:P	TB3	560	560	443	0.68	0.000	0.39
	Canopy N	SR	443	560	\	0.64	0.000	0.21
	Canopy N	SD	2190	705	\	0.55	0.000	0.24
	Canopy N	ND	443	560	\	0.63	0.000	0.22
	Canopy N	TB1	443	560	490	0.63	0.000	0.21
	Canopy N	TB2	705	490	2190	0.65	0.000	0.21
	Canopy N	TB3	2190	705	490	0.66	0.000	0.21
	Canopy P	SR	443	560	\	0.63	0.000	0.37
	Canopy P	SD	443	490	\	0.48	0.000	0.44
	Canopy P	ND	443	560	\	0.63	0.000	0.37
	Canopy P	TB1	443	560	490	0.64	0.000	0.36
Canopy P	TB2	443	560	490	0.62	0.000	0.37	
Canopy P	TB3	705	560	443	0.62	0.000	0.37	
OLCI	Canopy N:P	SR	412.5	560	\	0.67	0.000	0.40
	Canopy N:P	SD	490	510	\	0.50	0.000	0.50
	Canopy N:P	ND	412.5	560	\	0.67	0.000	0.40
	Canopy N:P	TB1	490	665	560	0.68	0.000	0.39

sensor	canopy trait	VI category	λ_1 (nm)	λ_2 (nm)	λ_3 (nm)	r^2	p -value	RRMSEcv
	Canopy N:P	TB2	560	442.5	400	0.67	0.000	0.40
	Canopy N:P	TB3	400	560	442.5	0.69	0.000	0.39
	Canopy N	SR	412.5	560	\	0.64	0.000	0.21
	Canopy N	SD	490	510	\	0.49	0.000	0.25
	Canopy N	ND	412.5	560	\	0.64	0.000	0.21
	Canopy N	TB1	490	620	560	0.64	0.000	0.21
	Canopy N	TB2	560	442.5	400	0.64	0.000	0.21
	Canopy N	TB3	400	560	442.5	0.66	0.000	0.21
	Canopy P	SR	490	510	\	0.63	0.000	0.37
	Canopy P	SD	490	510	\	0.49	0.000	0.44
	Canopy P	ND	490	510	\	0.63	0.000	0.37
	Canopy P	TB1	442.5	560	510	0.64	0.000	0.37
	Canopy P	TB2	510	490	412.5	0.64	0.000	0.36
	Canopy P	TB3	708.75	510	442.5	0.63	0.000	0.37
MODIS	Canopy N:P	SR	421.5	531	\	0.69	0.000	0.39
	Canopy N:P	SD	2130	555	\	0.50	0.000	0.50
	Canopy N:P	ND	421.5	531	\	0.68	0.000	0.39
	Canopy N:P	TB1	488	667	531	0.72	0.000	0.37
	Canopy N:P	TB2	531	412.5	421.5	0.70	0.000	0.38
	Canopy N:P	TB3	421.5	555	531	0.71	0.000	0.38
	Canopy N	SR	421.5	531	\	0.66	0.000	0.21
	Canopy N	SD	2130	555	\	0.54	0.000	0.24
	Canopy N	ND	421.5	531	\	0.65	0.000	0.21
	Canopy N	TB1	488	645	531	0.66	0.000	0.21
	Canopy N	TB2	531	421.5	555	0.67	0.000	0.21
	Canopy N	TB3	421.5	555	531	0.67	0.000	0.20
	Canopy P	SR	421.5	531	\	0.63	0.000	0.37
	Canopy P	SD	443	531	\	0.48	0.000	0.44
	Canopy P	ND	421.5	531	\	0.64	0.000	0.36
	Canopy P	TB1	531	551	421.5	0.64	0.000	0.36
	Canopy P	TB2	421.5	531	555	0.65	0.000	0.36
	Canopy P	TB3	551	469	443	0.63	0.000	0.37
OLI	Canopy N:P	SR	443	562	\	0.68	0.000	0.40
	Canopy N:P	SD	443	590	\	0.48	0.000	0.50
	Canopy N:P	ND	482	590	\	0.67	0.000	0.40
	Canopy N:P	TB1	443	590	562	0.68	0.000	0.39
	Canopy N:P	TB2	590	482	443	0.67	0.000	0.40
	Canopy N:P	TB3	562	562	443	0.69	0.000	0.39
	Canopy N	SR	443	562	\	0.64	0.000	0.21
	Canopy N	SD	2200	562	\	0.51	0.000	0.25
	Canopy N	ND	443	562	\	0.63	0.000	0.22
	Canopy N	TB1	482	590	562	0.64	0.000	0.21
	Canopy N	TB2	562	482	2200	0.64	0.000	0.21
	Canopy N	TB3	2200	562	482	0.66	0.000	0.21
	Canopy P	SR	443	562	\	0.63	0.000	0.37
	Canopy P	SD	443	562	\	0.46	0.000	0.45
	Canopy P	ND	443	562	\	0.64	0.000	0.37
	Canopy P	TB1	443	562	482	0.64	0.000	0.37
	Canopy P	TB2	562	443	482	0.63	0.000	0.37
	Canopy P	TB3	562	562	443	0.62	0.000	0.38

sensor	canopy trait	VI category	λ_1 (nm)	λ_2 (nm)	λ_3 (nm)	r^2	p -value	RRMSEcv
WorldView 4	Canopy N:P	SR	480	545	\	0.63	0.000	0.42
	Canopy N:P	SD	480	545	\	0.48	0.000	0.51
	Canopy N:P	ND	480	545	\	0.63	0.000	0.42
	Canopy N:P	TB1	480	672	545	0.67	0.000	0.40
	Canopy N:P	TB2	545	480	PAN	0.55	0.000	0.47
	Canopy N:P	TB3	480	480	545	0.64	0.000	0.42
	Canopy N	SR	480	545	\	0.61	0.000	0.22
	Canopy N	SD	480	545	\	0.48	0.000	0.26
	Canopy N	ND	480	545	\	0.60	0.000	0.22
	Canopy N	TB1	545	545	480	0.60	0.000	0.22
	Canopy N	TB2	545	480	PAN	0.56	0.000	0.23
	Canopy N	TB3	PAN	545	480	0.64	0.000	0.21
	Canopy P	SR	545	480	\	0.60	0.000	0.38
	Canopy P	SD	480	545	\	0.46	0.000	0.45
	Canopy P	ND	480	545	\	0.59	0.000	0.38
	Canopy P	TB1	480	480	545	0.60	0.000	0.38
	Canopy P	TB2	480	545	PAN	0.54	0.000	0.41
Canopy P	TB3	PAN	545	480	0.60	0.000	0.38	
RapidEye	Canopy N:P	SR	475	555	\	0.64	0.000	0.42
	Canopy N:P	SD	475	555	\	0.48	0.000	0.51
	Canopy N:P	ND	475	555	\	0.65	0.000	0.42
	Canopy N:P	TB1	475	658	555	0.67	0.000	0.40
	Canopy N:P	TB2	555	475	710	0.59	0.000	0.45
	Canopy N:P	TB3	710	555	475	0.66	0.000	0.41
	Canopy N	SR	475	555	\	0.61	0.000	0.22
	Canopy N	SD	475	555	\	0.47	0.000	0.26
	Canopy N	ND	475	555	\	0.61	0.000	0.22
	Canopy N	TB1	555	555	475	0.61	0.000	0.22
	Canopy N	TB2	555	475	710	0.60	0.000	0.22
	Canopy N	TB3	710	555	475	0.65	0.000	0.21
	Canopy P	SR	555	475	\	0.61	0.000	0.38
	Canopy P	SD	475	555	\	0.46	0.000	0.45
	Canopy P	ND	475	555	\	0.60	0.000	0.38
	Canopy P	TB1	475	475	555	0.60	0.000	0.38
	Canopy P	TB2	555	475	710	0.57	0.000	0.40
Canopy P	TB3	710	555	475	0.61	0.000	0.38	

Table A. 2.3. Relationship between existing vegetation indices (VIs), calculated with the spectra resampled to satellite sensors of four satellites, with canopy N:P, canopy N and canopy P. The satellite sensors included MSI, aboard Sentinel 2, OLCI, aboard Sentinel 3, MODIS aboard Terra-Aqua, and OLI, aboard Landsat 8. No existing VI was calculated for the RapidEye and WorldView 4 sensors because their band positions are too distant (> 100 nm) from the nominal VIs' wavelengths. Coefficient of determination (r^2), p-value and Relative Root Mean Squared Error of cross-validation (RRMSEcv). The model between the variables is either linear (L) or logarithmic (LOG). The VIs were classified into four categories: ND = Normalized difference; SR = simple ratio; TB = Three band; OVI = other vegetation index.

Sensor	VI	Canopy N:P				Canopy N				Canopy P			
		Model	r^2	p-value	RRMSEcv	Model	r^2	p-value	RRMSEcv	Model	r^2	p-value	RRMSEcv
MSI	TB (R498, R413, R442)	LOG	0.62	0.000	0.43	LOG	0.58	0.000	0.23	LOG	0.62	0.000	0.38
	TB (R434, R496, R401)	L	0.63	0.000	0.43	L	0.58	0.000	0.23	L	0.62	0.000	0.38
	ND (R735, R1285)	L	0.30	0.000	0.58	L	0.29	0.000	0.30	L	0.21	0.001	0.53
	TB (R705, R717, R491)	LOG	0.41	0.000	0.54	LOG	0.35	0.000	0.29	L	0.35	0.000	0.48
	TB (R760, R705, R445)	L	0.31	0.000	0.58	L	0.31	0.000	0.29	LOG	0.32	0.000	0.50
	SR (R990, R720)	LOG	0.27	0.000	0.60	LOG	0.29	0.000	0.30	LOG	0.26	0.000	0.52
	OVI (R670, R700, R740, R780) REP	LOG	0.29	0.000	0.59	LOG	0.33	0.000	0.29	LOG	0.32	0.000	0.50
OLCI	TB (R498, R413, R442)	LOG	0.40	0.000	0.54	LOG	0.35	0.000	0.28	LOG	0.43	0.000	0.46
	TB (R434, R496, R401)	L	0.33	0.000	0.57	L	0.28	0.000	0.30	L	0.38	0.000	0.48
	TB (R705, R717, R491)	L	0.39	0.000	0.54	LOG	0.36	0.000	0.28	L	0.35	0.000	0.48
	TB (R760, R705, R445)	LOG	0.30	0.000	0.58	LOG	0.32	0.000	0.29	LOG	0.32	0.000	0.50
	SR (R990, R720)	LOG	0.30	0.000	0.59	LOG	0.32	0.000	0.29	LOG	0.28	0.000	0.51
	OVI (R670, R700, R740, R780) REP	L	0.29	0.000	0.59	L	0.32	0.000	0.29	L	0.32	0.000	0.50
	MODIS	TB (R498, R413, R442)	LOG	0.36	0.000	0.56	LOG	0.31	0.000	0.29	LOG	0.40	0.000
TB (R434, R496, R401)	L	0.36	0.000	0.55	L	0.31	0.000	0.29	L	0.40	0.000	0.47	
ND (R1220, R710)	L	0.07	0.063	0.68	L	0.07	0.068	0.34	L	0.05	0.144	0.59	
ND (R735, R1285)	L	0.23	0.001	0.61	L	0.25	0.000	0.31	L	0.14	0.008	0.55	
TB (R705, R717, R491)	LOG	0.02	0.286	0.70	LOG	0.02	0.327	0.36	LOG	0.02	0.392	0.60	
TB (R760, R705, R445)	L	0.08	0.050	0.68	L	0.06	0.101	0.35	L	0.06	0.089	0.58	
SR (R990, R720)	LOG	0.15	0.007	0.65	LOG	0.21	0.001	0.31	LOG	0.11	0.022	0.57	
OVI (R670, R700, R740, R780) REP	LOG	0.26	0.000	0.60	LOG	0.29	0.000	0.30	LOG	0.27	0.000	0.51	
OLI	TB (R498, R413, R442)	LOG	0.55	0.000	0.47	LOG	0.50	0.000	0.25	LOG	0.55	0.000	0.41
	TB (R434, R496, R401)	L	0.55	0.000	0.47	L	0.50	0.000	0.25	L	0.55	0.000	0.41
	TB (R705, R717, R491)	LOG	0.27	0.000	0.60	LOG	0.21	0.001	0.32	LOG	0.19	0.002	0.54

2.8 References

- Bakker, M.A., Carreño-Rocabado, G., & Poorter, L. (2011). Leaf economics traits predict litter decomposition of tropical plants and differ among land use types. *Functional Ecology*, 25, 473-483. 10.1111/j.1365-2435.2010.01802.x
- Blackburn, G.A. (1998). Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *International Journal of Remote Sensing*, 19, 657-675. 10.1080/014311698215919
- Buschmann, C., & Nagel, E. (1993). In vivo spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. *International Journal of Remote Sensing*, 14, 711-722. 10.1080/01431169308904370
- Chen, J.M. (1996). Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Canadian Journal of Remote Sensing*, 22, 229-242. 10.1080/07038992.1996.10855178
- Cho, M.A., Ramoelo, A., Debba, P., Mutanga, O., Mathieu, R., van Deventer, H., & Ndlovu, N. (2013). Assessing the effects of subtropical forest fragmentation on leaf nitrogen distribution using remote sensing data. *Landscape Ecology*, 28, 1479-1491. doi:10.1007/s10980-013-9908-7
- Clevers, J.G.P.W., & Gitelson, A.A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and -3. *International Journal of Applied Earth Observation and Geoinformation*, 23, 344-351. doi:10.1016/j.jag.2012.10.008
- Corp, L.A., Middleton, E.M., Campbell, P.K.E., Huemmrich, K.F., Daughtry, C.S.T., Russ, A.L., & Cheng, Y.B. (2010). Spectral indices to monitor nitrogen-driven carbon uptake in field corn. *Journal of Applied Remote Sensing*, 4. 10.1117/1.3518455
- Dash, J., & Curran, P.J. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25, 5403-5413. doi:10.1080/0143116042000274015
- Evans, J.R. (1989). Photosynthesis and nitrogen relationships in leaves of C3 plants. *Oecologia*, 78, 9-19. doi:10.1007/BF00377192
- Ferwerda, J.G., Skidmore, A.K., & Mutanga, O. (2005). Nitrogen detection with hyperspectral normalized ratio indices across multiple plant species. *International Journal of Remote Sensing*, 26, 4083-4095. 10.1080/01431160500181044
- Fitzgerald, G., Rodriguez, D., & O'Leary, G. (2010). Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index – The canopy chlorophyll content index (CCCI). *Field Crops Research*, 116, 318-324. http://dx.doi.org/10.1016/j.fcr.2010.01.010
- Fujita, Y., Venterink, H.O., Van Bodegom, P.M., Douma, J.C., Heil, G.W., Hölzel, N., Jabłońska, E., Kotowski, W., Okruszko, T., Pawlikowski, P., De Ruiter, P.C., & Wassen, M.J. (2014). Low investment in sexual reproduction threatens plants adapted to phosphorus limitation. *Nature*, 505, 82-86. 10.1038/nature12733
- Gamon, J.A., Serrano, L., & Surfus, J.S. (1997). The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia*, 112, 492-501. 10.1007/s004420050337
- Gitelson, A.A., Kaufman, Y.J., & Merzlyak, M.N. (1996). Use of a green channel in remote sensing of global vegetation from EOS- MODIS. *Remote Sensing of Environment*, 58, 289-298. 10.1016/S0034-4257(96)00072-7
- Gitelson, A.A., Viña, A., Ciganda, V., Rundquist, D.C., & Arkebauer, T.J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, 32, 1-4. 10.1029/2005GL022688

- Gökkaya, K., Thomas, V., Noland, T., McCaughey, H., Morrison, I., & Treitz, P. (2015). Mapping continuous forest type variation by means of correlating remotely sensed metrics to canopy N: P ratio in a boreal mixedwood forest. *Applied Vegetation Science*, 18, 143-157. 10.1111/avsc.12122
- Green, D.S., Erickson, J.E., & Kruger, E.L. (2003). Foliar morphology and canopy nitrogen as predictors of light-use efficiency in terrestrial vegetation. *Agricultural and Forest Meteorology*, 115, 163-171. [http://dx.doi.org/10.1016/S0168-1923\(02\)00210-1](http://dx.doi.org/10.1016/S0168-1923(02)00210-1)
- Güsewell, S. (2004). N:P ratios in terrestrial plants: Variation and functional significance. *New Phytologist*, 164, 243-266. 10.1111/j.1469-8137.2004.01192.x
- Güsewell, S. (2005). High nitrogen: Phosphorus ratios reduce nutrient retention and second-year growth of wetland sedges. *New Phytologist*, 166, 537-550. 10.1111/j.1469-8137.2005.01320.x
- Güsewell, S., Koerselman, W., & Verhoeven, J.T.A. (2003). Biomass N:P ratios as indicators of nutrient limitation for plant populations in wetlands. *Ecological Applications*, 13, 372-384. 10.1890/1051-0761(2003)013[0372:BNRAIO]2.0.CO;2
- Hansen, P.M., & Schjoerring, J.K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86, 542-553. doi:10.1016/S0034-4257(03)00131-7
- Homolová, L., Malenovský, Z., Clevers, J.G.P.W., García-Santos, G., & Schaepman, M.E. (2013). Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity*, 15, 1-16. 10.1016/j.ecocom.2013.06.003
- Horler, D.N.H., Dockray, M., & Barber, J. (1983). The red edge of plant leaf reflectance. *International Journal of Remote Sensing*, 4, 273-288. 10.1080/01431168308948546
- Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295-309. 10.1016/0034-4257(88)90106-X
- Huete, A.R., & Jackson, R.D. (1987). Suitability of spectral indices for evaluating vegetation characteristics on arid rangelands. *Remote Sensing of Environment*, 23, 213-218. [http://dx.doi.org/10.1016/0034-4257\(87\)90038-1](http://dx.doi.org/10.1016/0034-4257(87)90038-1)
- Jordan, C.F. (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology*, 50, 663-666. 10.2307/1936256
- Kawamura, K., Mackay, A.D., Tuohy, M.P., Betteridge, K., Sanches, I.D., & Inoue, Y. (2011). Potential for spectral indices to remotely sense phosphorus and potassium content of legume-based pasture as a means of assessing soil phosphorus and potassium fertility status. *International Journal of Remote Sensing*, 32, 103-124. 10.1080/01431160903439908
- Kergoat, L., Lafont, S., Arneth, A., Le Dantec, V., & Saugier, B. (2008). Nitrogen controls plant canopy light-use efficiency in temperate and boreal ecosystems. *Journal of Geophysical Research: Biogeosciences*, 113. doi:10.1029/2007JG000676
- 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. 10.2307/2404783
- Kumar, L., Schmidt, K., Dury, S., & Skidmore, A. (2006). Imaging Spectrometry and Vegetation Science. In F.D.v.d. Meer, & S.M. de Jong (Eds.), *Imaging Spectrometry: Basic Principles and Prospective Applications* (pp. 111-155). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-0-306-47578-8_5. ISBN: 978-0-306-47578-8
- le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailier, J.-Y., Bréda, N., Genet, H., Davi, H., & Dufrêne, E. (2008). Calibration and validation of hyperspectral indices for the estimation of broad-

- leaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment*, 112, 3846-3864. <http://dx.doi.org/10.1016/j.rse.2008.06.005>
- LeBauer, D.S., & Treseder, K.K. (2008). Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, 89, 371-379. 10.1890/06-2057.1
- Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., Liu, Y., Liu, B., Ustin, S.L., & Chen, X. (2014). Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Research*, 157, 111-123. doi:10.1016/j.fcr.2013.12.018
- Li, L.J., Zeng, D.H., Yu, Z.Y., Fan, Z.P., Mao, R., & Peri, P.L. (2011). Foliar N/P ratio and nutrient limitation to vegetation growth on Keerqin sandy grassland of North-east China. *Grass and Forage Science*, 66, 237-242. 10.1111/j.1365-2494.2011.00781.x
- Loozen, Y., Rebel, K.T., Karssenbergh, D., Wassen, M.J., Sardans, J., Peñuelas, J., & de Jong, S.M. (2018). Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index. *Biogeosciences*, 15, 2723-2742. 10.5194/bg-15-2723-2018
- Mahajan, G.R., Sahoo, R.N., Pandey, R.N., Gupta, V.K., & Kumar, D. (2014). Using hyperspectral remote sensing techniques to monitor nitrogen, phosphorus, sulphur and potassium in wheat (*Triticum aestivum* L.). *Precision Agriculture*, 15, 499-522. 10.1007/s11119-014-9348-7
- 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. doi:10.2111/04-17.1
- Olde Venterink, H., Wassen, M.J., Verkoost, A.W.M., & De Ruiter, P.C. (2003). Species richness-productivity patterns differ between N-, P-, and K-limited wetlands. *Ecology*, 84, 2191-2199. 10.1890/01-0639
- Pacheco-Labrador, J., González-Cascón, R., Pilar Martín, M., & Riaño, D. (2014). Understanding the optical responses of leaf nitrogen in mediterranean holm oak (*Quercus ilex*) using field spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, 26, 105-118. doi:10.1016/j.jag.2013.05.013
- Peñuelas, J., Baret, F., & Filella, I. (1995). Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica*, 31, 221-230
- Pimstein, A., Karnieli, A., Bansal, S.K., & Bonfil, D.J. (2011). Exploring remotely sensed technologies for monitoring wheat potassium and phosphorus using field spectroscopy. *Field Crops Research*, 121, 125-135. 10.1016/j.fcr.2010.12.001
- R Development Core Team (2014). *R: A Language and Environment for Statistical Computing* [Computer program]. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>
- Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., & Heitkönig, I.M.A. (2012). Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor. *International Journal of Applied Earth Observation and Geoinformation*, 19, 151-162. doi:10.1016/j.jag.2012.05.009
- Ramoelo, A., Skidmore, A.K., Schlerf, M., Heitkönig, I.M.A., Mathieu, R., & Cho, M.A. (2013). Savanna grass nitrogen to phosphorous ratio estimation using field spectroscopy and the potential for estimation with imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, 23, 334-343. 10.1016/j.jag.2012.10.008
- Reich, P.B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society B: Biological Sciences*, 279, 2128-2134. doi:10.1098/rspb.2011.2270

- Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Volin, J.C., & Bowman, W.D. (1999). Generality of leaf trait relationships: A test across six biomes. *Ecology*, *80*, 1955-1969. doi:10.2307/176671
- Roeling, I.S., Ozinga, W.A., van Dijk, J., Eppinga, M.B., & Wassen, M.J. (2018). Plant species occurrence patterns in Eurasian grasslands reflect adaptation to nutrient ratios. *Oecologia*, *186*, 1055-1067. 10.1007/s00442-018-4086-6
- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, *55*, 95-107. [http://dx.doi.org/10.1016/0034-4257\(95\)00186-7](http://dx.doi.org/10.1016/0034-4257(95)00186-7)
- Roujean, J.-L., & Breon, F.-M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, *51*, 375-384. [http://dx.doi.org/10.1016/0034-4257\(94\)00114-3](http://dx.doi.org/10.1016/0034-4257(94)00114-3)
- Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, *25*, 47-54. doi:10.1016/j.jag.2013.04.003
- Serrano, L., Peñuelas, J., & Ustin, S.L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*, 355-364. doi:10.1016/S0034-4257(02)00011-1
- 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. [http://dx.doi.org/10.1016/S0034-4257\(02\)00010-X](http://dx.doi.org/10.1016/S0034-4257(02)00010-X)
- 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. 10.1046/j.1365-2664.2003.00820.x
- Thenkabail, P.S., Smith, R.B., & De Pauw, E. (2002). Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. *Photogrammetric Engineering and Remote Sensing*, *68*, 607-621
- Tian, Y.C., Yao, X., Yang, J., Cao, W.X., Hannaway, D.B., & Zhu, Y. (2011). Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance. *Field Crops Research*, *120*, 299-310. doi:10.1016/j.fcr.2010.11.002
- Vitousek, P.M., & Howarth, R.W. (1991). Nitrogen limitation on land and in the sea: How can it occur? *Biogeochemistry*, *13*, 87-115. 10.1007/BF00002772
- Walker, A.P., Beckerman, A.P., Gu, L., Kattge, J., Cernusak, L.A., Domingues, T.F., Scales, J.C., Wohlfahrt, G., Wullschlegel, S.D., & Woodward, F.I. (2014). The relationship of leaf photosynthetic traits – V_{cmax} and J_{max} – to leaf nitrogen, leaf phosphorus, and specific leaf area: a meta-analysis and modeling study. *Ecology and Evolution*, *4*, 3218-3235. 10.1002/ece3.1173
- Wang, J., Shi, T., Liu, H., & Wu, G. (2016a). Successive projections algorithm-based three-band vegetation index for foliar phosphorus estimation. *Ecological Indicators*, *67*, 12-20. 10.1016/j.ecolind.2016.02.033
- Wang, J., Wang, T., Skidmore, A.K., Shi, T., & Wu, G. (2015). Evaluating different methods for grass nutrient estimation from canopy hyperspectral reflectance. *Remote Sensing*, *7*, 5901-5917. 10.3390/rs70505901
- Wang, W., Yao, X., Yao, X., Tian, Y., Liu, X., Ni, J., Cao, W., & Zhu, Y. (2012). Estimating leaf nitrogen concentration with three-band vegetation indices in rice and wheat. *Field Crops Research*, *129*, 90-98. <http://dx.doi.org/10.1016/j.fcr.2012.01.014>

- Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A.K., Jones, S., Suarez, L., Woodgate, W., Heiden, U., Heurich, M., & Hearne, J. (2016b). Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sensing*, 8. doi:10.3390/rs8060491
- Wassen, M.J., de Boer, H.J., Fleischer, K., Rebel, K.T., & Dekker, S.C. (2013). Vegetation-mediated feedback in water, carbon, nitrogen and phosphorus cycles. *Landscape Ecology*, 28, 599-614. 10.1007/s10980-012-9843-z
- Wassen, M.J., Venterink, H.O., Lapshina, E.D., & Tanneberger, F. (2005). Endangered plants persist under phosphorus limitation. *Nature*, 437, 547-550. 10.1038/nature03950
- Wassen, M.J., Venterink, H.O.G.M., & de Swart, E.O.A.M. (1995). Nutrient concentrations in mire vegetation as a measure of nutrient limitation in mire ecosystems. *Journal of Vegetation Science*, 6, 5-16. doi:10.2307/3236250
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., & Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, 428, 821-827. doi:10.1038/nature02403
- Xue, L., Cao, W., Luo, W., Dai, T., & Zhu, Y. (2004). Monitoring Leaf Nitrogen Status in Rice with Canopy Spectral Reflectance. *Agronomy Journal*, 96, 135-142. 10.2134/agronj2004.0135
- Yao, X., Zhu, Y., Tian, Y., Feng, W., & Cao, W. (2010). Exploring hyperspectral bands and estimation indices for leaf nitrogen accumulation in wheat. *International Journal of Applied Earth Observation and Geoinformation*, 12, 89-100. <http://dx.doi.org/10.1016/j.jag.2009.11.008>
- Zarco-Tejada, P.J., Berjón, A., López-Lozano, R., Miller, J.R., Martín, P., Cachorro, V., González, M.R., & De Frutos, A. (2005). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment*, 99, 271-287. 10.1016/j.rse.2005.09.002
- Zhu, Y., Tian, Y., Yao, X., Liu, X., & Cao, W. (2007a). Analysis of common canopy reflectance spectra for indicating leaf nitrogen concentrations in wheat and rice. *Plant Production Science*, 10, 400-411. 10.1626/ppp.10.400
- Zhu, Y., Zhou, D., Yao, X., Tian, Y., & Cao, W. (2007b). Quantitative relationships of leaf nitrogen status to canopy spectral reflectance in rice. *Australian Journal of Agricultural Research*, 58, 1077-1085. 10.1071/AR06413

Chapter 3

Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index



Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index²

Abstract. Canopy nitrogen (N) concentration and content are linked to several vegetation processes. Therefore, canopy N concentration is a state variable in global vegetation models with coupled carbon (C) and N cycles. While there are ample C data available to constrain the models, widespread N data are lacking. Remotely sensed vegetation indices have been used to detect canopy N concentration and canopy N content at the local scale in grasslands and forests. Vegetation indices could be a valuable tool to detect canopy N concentration and canopy N content at larger scale. In this paper we conducted a regional case-study analysis to investigate the relationship between the Medium Resolution Imaging Spectrometer (MERIS) Terrestrial Chlorophyll Index (MTCI) time series from European Space Agency (ESA) Envisat satellite at 1 km spatial resolution and both canopy N concentration (%N) and canopy N content (N g m⁻² of ground area) from a Mediterranean forest inventory in the region of Catalonia, in the northeast of Spain. The relationships between the datasets were studied after resampling both datasets to lower spatial resolutions (20 km, 15 km, 10 km and 5 km) and at the original spatial resolution of 1 km. The results at higher spatial resolution (1 km) yielded significant log-linear relationships between MTCI and both canopy N concentration and content: $r^2 = 0.32$ and $r^2 = 0.17$, respectively. We also investigated these relationships per plant functional type. While the relationship between MTCI and canopy N concentration was strongest for deciduous broadleaf and mixed plots ($r^2 = 0.24$ and $r^2 = 0.44$, respectively), the relationship between MTCI and canopy N content was strongest for evergreen needleleaf trees ($r^2 = 0.19$). At the species level, canopy N concentration was strongly related to MTCI for European beech plots ($r^2 = 0.69$). These results present a new perspective on the application of MTCI time series for canopy N detection.

Keywords: vegetation index, MERIS, foliar nitrogen concentration, foliar nitrogen content, plant functional types, Mediterranean forest, remote sensing

2 This chapter is based on: Loozen, Y., Rebel, K.T., Karssenber, D., Wassen, M.J., Sardans, J., Peñuelas, J., & de Jong, S.M. (2018). Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index. *Biogeosciences*, 15, 2723-2742. <https://doi.org/10.5194/bg-15-2723-2018>

3.1 Introduction

Canopy nitrogen (N) concentration is an essential state variable in regional (Ollinger and Smith, 2005) and global vegetation models including both the carbon (C) and the N cycles (such as Zaehle and Friend 2010; Smith 2014). This variable has been linked to several vegetation traits and processes at the leaf and canopy levels. At the leaf level, leaf N concentration, which represents the leaf N status expressed as a percentage of leaf dry matter (%N, $\text{N g } 100\text{g}^{-1}\text{DM}$), has been related to photosynthetic capacity (Evans 1989; Reich et al. 1995; Reich et al. 1997; Reich et al. 1999; Wright et al. 2004), specific leaf area, leaf life span (Reich et al. 1999; Wright et al. 2004) and light use efficiency (Kergoat et al. 2008). Leaf N concentration expressed on a leaf area basis, also called leaf N content (N g m^{-2}), has also been linked with chlorophyll content, RuBisCO content (Evans 1989) and photosynthetic capacity (Evans 1989; Reich et al. 1995). At stand scale, canopy N concentration, which represents the leaf N concentration averaged over the stand canopy, has also been found to correlate with above-ground net primary productivity (NPP) (Reich 2012), while canopy N content has been linked with the canopy light use efficiency (Green et al. 2003).

Given their links to many vegetation processes, leaf and canopy N variables could be used to constrain N cycle modules in global vegetation models. At the global scale, ample data are available to constrain models for the C cycle; however, data to constrain the N cycle are limited. Currently, canopy N data are not widely available and canopy N sampling campaigns are time consuming and thus expensive tasks. Moreover, upscaling from local sampling campaign measurements represents an additional limitation. From this perspective, local, regional or even global remotely sensed canopy N estimates will be a valuable addition, enabling us to collect information in a less time-intensive and expensive manner than traditional on-field sampling campaigns. Such near-global canopy N estimates will be beneficial as input in global vegetation models or to calibrate and validate these models.

Remote detection of foliage N status has been extensively studied at the leaf scale (Hansen and Schjoerring 2003; Ferwerda et al. 2005; Li et al. 2014) and few studies have investigated the processes underlying the relationships between vegetation indices and foliar N (Pacheco-Labrador et al. 2014). Detection of foliage N status with vegetation indices is attributed to the strong link between foliar N and chlorophyll content (Schlemmer et al. 2013) and is often based on the near-infrared (NIR) and red-edge regions of the spectrum, hence similar to the ones used for chlorophyll detection (Filella and Penuelas 1994; Dash and Curran 2004; Clevers and Gitelson 2013). At canopy level, however, spectral reflectance is a complex function of vegetation cover, plant activity, water content, illumination angle, viewing angle and atmospheric composition (Kumar et al. 2006) and it is not straightforward to disentangle the influence of nitrogen from other contributions in the spectra. It is thus not clear how the relationships observed at the leaf level translate at the canopy level. The mechanisms possibly modifying the remote sensing of foliage N status at the canopy scale are still not clearly understood (Ollinger 2011). High correlation between canopy N and both NIR reflectance and albedo has been reported in boreal forests (Ollinger et al. 2008). However, the mechanism behind these findings is still controversial. Knyazikhin et al. (2013) argued that the observed correlation solely resulted from canopy structural differences between broad and needleleaf forests and was thus spurious. Other authors, although agreeing that canopy structure was

a confounding factor to account for, stated that the NIR-canopy N relationship was not necessarily spurious and stemmed from an association between canopy N and structural traits (Ollinger et al. 2013; Townsend et al. 2013). Canopy traits are interrelated (Wright et al. 2004) and have been known to covary due to evolutionary convergence, as stated by Ollinger (2011).

Different remote sensing techniques have been applied to detect canopy N in terrestrial vegetation. Imaging spectrometry has proven efficient in improving N sensing capabilities at the local scale. Imaging spectrometry images are acquired from either airborne or spaceborne sensors and are analysed with different methods, including partial least squares regression (PLS), continuum removal, spectral unmixing or vegetation indices (Smith et al. 2003; Ollinger et al. 2008; Huber et al. 2008; Martin et al. 2008; Schlerf et al. 2010; Wang et al. 2016). Among other techniques, ratios or normalised differences of reflectance bands in the red and NIR regions of the spectrum, the so-called vegetation indices (VIs) (Glenn et al. 2008), are one of the most straightforward methods for canopy N detection. Combined with in situ hyperspectral devices, vegetation indices have been extensively used for leaf or canopy N detection in agricultural systems (Peñuelas et al. 1994; Filella et al. 1995; Hansen and Schjoerring 2003; Tian et al. 2011; Schlemmer et al. 2013; Li et al. 2014). Vegetation indices have also been applied to airborne or spaceborne acquired imagery in natural environments (Ramoelo et al. 2012; Wang et al. 2016).

A particular vegetation index, the Medium Resolution Imaging Spectrometer (MERIS) Terrestrial Chlorophyll Index (MTCI) has been proposed for detecting canopy N (Clevers and Gitelson 2013). MTCI was originally computed from three reflectance bands from MERIS aboard the European Space Agency (ESA) Envisat satellite at a spatial resolution of 1 km. However, it can also be obtained from other sensors' reflectance data and a similar product will be available from the ESA Sentinel-2 satellite mission (Drusch et al. 2012). It was first developed to estimate chlorophyll content (Dash and Curran 2004, 2007). Regarding canopy N detection, most studies were carried out in agricultural crops using MTCI values computed from in situ hyperspectral reflectance data (Tian et al. 2011; Clevers and Gitelson 2013; Li et al. 2014). A few were directed towards sensing N concentration in natural environments using airborne data, e.g. in temperate forests (Wang et al. 2016), or spaceborne data, for example in grasslands (Ramoelo et al. 2012; Ullah et al. 2012) or subtropical forests (Cho et al. 2013).

In this context, there are several knowledge gaps that we would like to address in this paper. First, although 1 km spatial resolution spaceborne MTCI time series are available from the ESA, MTCI has mainly been employed to detect canopy N in agricultural applications with in situ devices and rarely in a broader range of natural ecosystems and scales using spaceborne data. Due to its almost global coverage, MTCI time series could be applied to estimate canopy N over a larger spatial extent. Moreover, Mediterranean forests have specific functional characteristic due to their great forest ecosystems diversity, influenced by contrasting climatic and topographic conditions, and their high tree species richness (Vilà-Cabrera et al. 2018). However, to our knowledge, limited research has been conducted to sense canopy N in Mediterranean ecosystems (Serrano et al. 2002) and even more so in Mediterranean forests. The relationship between MTCI and both N concentration ($N_{[\%]}$, %N) and canopy N content ($N_{[\text{area}]}$, g m^{-2}) has been studied separately (Clevers and Gitelson 2013; Wang et al. 2016), but very few analyses (Mirik et al. 2005; Ullah et al. 2012) have compared the ability to detect canopy N concentration and canopy N content simultaneously, especially in forest ecosystems.

The objective of our study is thus to investigate the relationship between the spaceborne MTCI remote sensing product and canopy N in Mediterranean forests at the regional scale. More specifically, the relationships between MTCI and both canopy N concentration and canopy N content are investigated and compared. We then also examine these relationships per plant functional type (PFT) and at the species level.

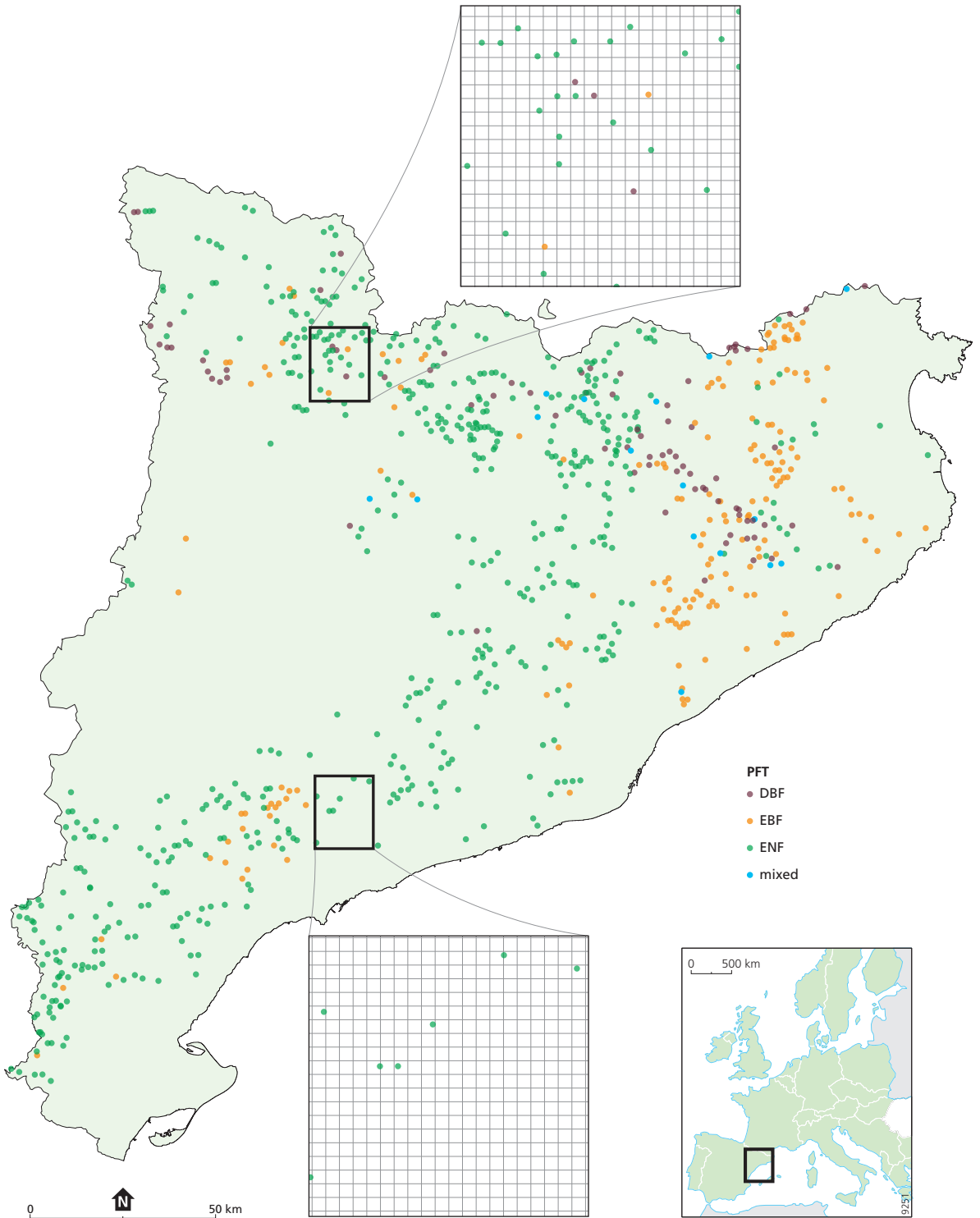
Remote sensing of canopy N is often limited to local-scale studies due to the spatial restrictions associated with N data acquisition in the field and treatment of high spatial resolution remote sensing imagery with limited spatial coverage (Lepine et al. 2016). Our case study exploits the broadly and readily available MTCI time series at 1 km spatial resolution from the ESA Envisat mission and combines it with canopy N data, both concentration and content, from 846 forest plots measured between 1988 and 2001 by the Catalanian National Forest Inventory (Gracia et al. 2004). First, we develop a methodology to overcome the time discrepancy between our two sets of data. Next, both datasets are resampled to the same lower spatial resolutions, i.e. 5 km, 10 km, 15 km and 20 km, in order to overcome the initial spatial discrepancy between MTCI spatial resolution (1 km) and the size of the forest plots (6 m). Subsequently, we analyse the relationship between MTCI and both canopy N concentration and canopy N content variables, both at the resampled and initial spatial resolutions. The relationships at the initial spatial resolution are then stratified according to the PFT of the plots. The results are presented and discussed. Finally, we address the implications for future research and draw a conclusion.

3.2 Material and methods

3.2.1 Study area

Our study area corresponds to the region of Catalonia (Figure 3.1) which is located in northeastern Spain and has a spatial extent of 32,114 km² (Sardans et al. 2011). While the region is characterised by a Mediterranean climate, the presence of the Pyrenees to the northwest and the Mediterranean Sea to the east creates contrasting climate conditions with an altitudinal gradient from north to south and a continental gradient from west to east. Following this pattern, the mean annual temperature varies from 1 °C in the north to 17 °C in the south (Sardans et al. 2011). While mean annual precipitation (MAP) is 1400 mm in the Pyrenees, in the south, the MAP is lower than 350 mm (Sardans et al. 2011), leading to seasonal drought (Lana and Burgueño, 1998) and fires (González and Pukkala 2007), which impact the vegetation (Liu et al. 2015).

Figure 3.1 (next page) Map showing the forest plots (n = 846) location in the region of Catalonia, northeastern Spain. Two zoom windows are included showing the density of the plots, one with high density and one with low density, relative to the MTCI 1 km pixel grid. DBF is deciduous broadleaf forest, EBF is evergreen broadleaf forest, ENF is evergreen needleleaf forest, mixed is mixed forest.



3.2.2 Data collection

3.2.2.1 Canopy N data

The canopy N data used in this research were collected by the Ecological and Forestry Applications Research Centre (CREAF), Universitat Autònoma de Barcelona. The data included 2300 closed canopy forest plots sampled between 1988 and 2001 by the Catalanian National Forest Inventory (Gracia et al. 2004).

The forest plots (Figure 3.1) had a minimum diameter of 6 m, which varied depending on the tree density in order to include between 15 and 25 trees with a diameter at breast height (DBH) of at least 5 cm. The DBH was recorded for all the trees present on the plot with a DBH of minimum 5 cm. The plots were investigated for canopy N concentration ($N_{[\%]}$, %N) defined as grams of N per 100 g of leaf dry matter. The leaf samples were collected from the upper central part of the crown using extensible loppers. All foliar cohorts present in the canopy were included in the leaf sample. Each leaf sample was constituted by the leaves of at least three different trees of the dominant tree species in the canopy. The species dominance was determined by the tallest individual. The proportion of 96% of the plots included in this analysis was monospecific (Sardans et al. 2011). A total of 4% of the plots ($n = 30$) had two codominant species. For these plots, two leaf samples were collected, one for each of the codominant species found on the plots.

The leaf samples were dried and then ground using a Braun Mikro-Dismembrator U (B. Braun Biotech International, Melsungen, Germany). They were analysed for foliar N concentration using the combustion technique coupled to gas chromatography using a Thermo Electron gas chromatograph (model NA 2100, CE Instruments, Thermo Electron, Milan, Italy) (Gracia et al. 2004). To scale from leaf to canopy level, we used the leaf N concentration averaged over three individuals as the plot level value (Schlerf et al. 2010). We did not weight the average by species abundance (Smith and Martin 2001) as only 4% of the plots had two different species.

Along with the canopy $N_{[\%]}$ data, we used foliar biomass data (grams of dry matter per square metre of ground area, $g\ m^{-2}$) acquired during the same forest inventory ($n = 2286$). The foliar biomass data were obtained for each plot from allometric equations relating the diameter at breast height to the leaves dry weight. The allometric equations were species specific (Sardans et al. (2015), Table A. 3.1). The foliar biomass data were used to calculate canopy N content ($N_{[area]}$, grams of N per square metre of ground, $g\ m^{-2}$) for each plot following Equation 3.1:

$$\text{canopy } N_{[area]} = \frac{\text{canopy } N_{[\%]} * f_{biom}}{100}, \quad (\text{Equation 3.1})$$

where *canopy* $N_{[area]}$ is the canopy N content ($g\ m^{-2}$), *canopy* $N_{[\%]}$ is the canopy N concentration (%N) and f_{biom} is the foliar biomass ($g\ m^{-2}$).

For the plots with two codominant species, the concentration measurements were done separately. The obtained foliar N concentration and biomass values were then averaged to obtain a single canopy $N_{[\%]}$ and $N_{[area]}$ value for each plot with two codominant species. Among these 30 plots with codominant species, 16 plots had codominant species from

Table 3.1. Descriptive analysis of canopy nitrogen (N) concentration (N_{fol} , $\text{g } 100\text{g}^{-1}$), foliar biomass (g m^{-2}) and canopy N content (N_{areal} , g m^{-2}) by tree species. PFT is plant functional type, DBF is deciduous broadleaf forest, EBF is evergreen broadleaf forest, ENF is evergreen needleleaf forest, mixed is mixed forest, min is minimum, max is maximum, mean is average, sd is standard deviation. ^a Codominant plots refer to the plots where two tree species were dominant in the canopy. ^b Foliar biomass data were lacking for five of the plots. Foliar biomass and canopy N content statistics are thus measured on a restricted number of plots.

Species	PFT	No. of plots	Abundance (% of total no. of plots)	Canopy N _{fol} (g 100g ⁻¹)			Foliar biomass (g m ⁻²)			Canopy N _{areal} (g m ⁻²)					
				min	max	mean	sd	min	max	mean	sd	min	max	mean	sd
<i>Castanea sativa</i>	DBF	14	1.7	1.62	2.81	2.08	0.36	18.13	425.90	203.46	123.49	0.40	11.99	4.25	2.89
<i>Fagus sylvatica</i>	DBF	15	1.8	1.22	3.13	2.28	0.61	49.94	279.86	173.54	68.70	1.21	7.40	3.96	1.95
<i>Pinus halepensis</i>	ENF	240	28.4	0.56	1.57	0.90	0.19	9.58	827.80	197.23	145.54	0.09	7.29	1.77	1.33
<i>Pinus nigra</i>	ENF	37	4.4	0.56	1.28	0.89	0.19	32.25	923.98	294.29	224.32	0.23	8.87	2.67	2.18
<i>Pinus pinaster</i>	ENF	5	0.6	0.82	1.08	0.93	0.13	271.75	718.87	501.67	211.53	2.30	7.69	4.75	2.25
<i>Pinus pinea</i>	ENF	5	0.6	0.75	1.06	0.95	0.14	103.28	275.50	179.74	66.80	1.08	2.91	1.71	0.75
<i>Pinus sylvestris</i>	ENF	198	23.4	0.67	2.14	1.11	0.20	10.48	828.63	326.44	181.20	0.10	12.86	3.65	2.22
<i>Pinus uncinata</i>	ENF	69	8.2	0.46	1.33	0.87	0.19	183.59	1744.50	687.22	345.21	1.41	16.97	5.92	3.25
<i>Quercus canariensis</i>	DBF	3	0.4	1.97	2.78	2.25	0.46	122.11	197.85	160.32	37.87	2.41	5.51	3.71	1.61
<i>Quercus faginea</i>	DBF	4	0.5	1.49	2.11	1.82	0.31	10.34	419.14	233.47	187.01	0.17	8.83	4.64	4.09
<i>Quercus humilis</i>	DBF	9	1.1	1.53	3.11	2.41	0.42	56.12	337.33	142.65	92.11	1.21	8.64	3.33	2.19
<i>Quercus cerritoides</i>	DBF	17	2.0	1.44	2.80	2.07	0.37	12.97	834.68	262.24	237.49	0.29	15.42	5.06	4.31
<i>Quercus ilex</i>	EBF	160	18.9	0.81	2.87	1.32	0.26	16.63	1033.31	378.23	238.61	0.22	16.61	4.95	3.23
<i>Quercus petraea</i>	DBF	17	2.0	1.37	2.70	2.21	0.41	20.45	741.42	279.96	229.78	0.32	15.37	5.98	4.66
<i>Quercus suber</i>	EBF	23	2.7	1.25	2.08	1.55	0.21	26.26	219.05	110.49	55.65	0.40	4.34	1.72	0.96
Codominant ^c	mixed	30 (25) ^b	3.5	0.92	2.54	1.45	0.41	23.45	342.58	153.70	77.39	0.33	5.74	2.06	1.02

different PFTs. Their PFT is thus labelled as mixed while the plots with several codominant species from the same PFT are labelled according to their PFTs.

Catalonian forests include both deciduous and evergreen broadleaf as well as evergreen needleleaf tree species. These three PFTs are referred to as deciduous broadleaf forest (DBF), evergreen broadleaf forest (EBF) and evergreen needleleaf forest (ENF), respectively. The main tree species are *Pinus halepensis* Mill., *Pinus sylvestris* L., *Quercus ilex* L., *Pinus uncinata* Ramond ex DC., *Pinus nigra* J.F. Arnold, *Quercus suber* L., *Quercus cerrrioides* Willk. & Costa., *Quercus petraea* Liebl. and *Fagus sylvatica* L. These species accounted for 92% of the sampled forest plots. The 15 tree species included in this analysis are listed in Table 3.1. Plots with a rare dominant tree species, i.e. species that were detected in only a single plot, were excluded from the analysis. This applied to plots with these dominant species: *Abies alba* Mill., *Fraxinus augustifolia* Vahl, *Fraxinus excelsior* L., *Pinus radiata* D. Don, *Populus nigra* L., *Populus tremula* L. and *Quercus robur* L.

3.2.2.2 MTCI product

MTCI was first developed to estimate chlorophyll content in canopies. MTCI is sensitive to high chlorophyll content, while presenting low sensitivity to soil brightness (Curran and Dash 2005). Its calculation, presented in Equation 3.2, is based on three reflectance bands, located around the red-edge point (REP) (Dash and Curran 2004):

$$MTCI = \frac{R_{band10} - R_{band9}}{R_{band9} - R_{band8}} = \frac{R_{753.75} - R_{708.75}}{R_{708.75} - R_{681.75}}, \quad (\text{Equation 3.2})$$

where R_{band8} , R_{band9} and R_{band10} represent the eighth, ninth, and tenth bands of MERIS, respectively. Following MERIS standard bands settings, the centres of the bands were located at 681.25 nm, 708.75 nm and 753.75 nm on the electromagnetic spectrum.

While the ESA Envisat satellite mission producing MERIS data came to an end in 2012, MERIS products and MTCI in particular are still relevant because the new ESA Sentinel-2 and Sentinel-3 satellite missions have improved band settings compared to those of MERIS. MTCI can be calculated from Sentinel-2 reflectance data with increased spatial resolution to 20 m (Drusch et al. 2012). The Sentinel-3 mission also releases a level-2 chlorophyll product, the Ocean and Land Colour Instrument (OLCI) Terrestrial Chlorophyll Index (OTCI), the calculation of which is directly based on MTCI. OTCI continues the time series already available for MTCI (Dash and Vuolo 2010; Vuolo et al. 2012). In this study, we put emphasis on Envisat-MERIS as our field data are closer to the MERIS acquisition period.

MTCI level-3 imagery was obtained from the Natural Environment Research Council (NERC) Earth Observation Data Centre (NEODC, 2015) for the region of Catalonia between 2002 and 2012. The original data were provided by the European Space Agency and then processed by Airbus Defence and Space. The original MERIS reflectance images, following Envisat specifications, have a revisit time of 3 days and a spatial resolution of 300 m. Compared to the original reflectance images, the MTCI-processed imagery has been corrected for atmospheric influences and cloud cover (Curran and Dash 2005), and is available as an either weekly or monthly averaged product almost globally (Curran et al. 2007). The spatial resolution of the processed data is approximately 1 km. As there is no temporally averaged

product available at full resolution, we chose to carry out this analysis with the MTCI monthly averaged processed imagery. This was done to decrease the uncertainty resulting from the use of single daily reflectance values. One MTCI monthly averaged imagery product covering the entire study area was obtained for every month between June 2002 and March 2012, except for October 2003, when no valid product was available.

3.2.3 Data handling

3.2.3.1 Methodology to link canopy N data to MTCI values

There is a discrepancy between the timing of the ground truth sampling and the satellite image acquisition period. While the plot sampling campaigns were carried out between 1988 and 2001, the Envisat satellite mission was launched in 2002 and ended in 2012. To overcome the discrepancy, MTCI images were averaged by month over the 10 years of the satellite mission period. This process yielded 12 MTCI averaged images, one for each month. The averaged MTCI images were then linked to the forest plots based on the forest plot coordinates and sampling month, as the exact sampling date was known for each plot. The period between 1 June and 31 of October was determined to be the growing season after a pre-analysis, where we studied yearly temporal variation of MTCI in several locations and forest types in Catalonia. This extended period was chosen to encompass the different vegetation phenology types corresponding to the contrasted climate conditions in this region. The forest plots sampled outside of the growing season were excluded from the analysis. The interannual variation of canopy $N_{[\%]}$ data was analysed for each month included in the analysis to ensure that the ground data could be related to MTCI data (Figure A. 3.1). The GlobCover 2009 land cover map was used to exclude forest plots for which the dominant vegetation type of the MTCI pixel did not correspond to natural vegetation. The GlobCover map was created by ESA using MERIS reflectance data from 2009 (Bontemps et al. 2011). The GlobCover map was downloaded from the ESA data user elements website (ESA, 2010). This map comprises 22 land cover classes and has a spatial resolution of 300 m. Using this map, we excluded forest plots that had undergone a land cover change since the sampling period and did not have a natural vegetation cover any more at the time of remote sensing image acquisition. To do so, the land cover map was first resampled to a spatial resolution of 1 km to be in accordance with MTCI spatial resolution. The resampling was done using the majority option, which ensured that the resampled land cover type was the most occurring land cover type in the MTCI pixel. Resampling the land cover map was done to exclude the plots located on heterogeneous MTCI pixels, i.e. pixels where the natural vegetation was not the dominant land cover type. Then, the plots located on land area classified as either rain-fed cropland, mosaic between croplands and natural vegetation, sparse vegetation or artificial surfaces were excluded from the analysis.

3.2.3.2 Relationship between MTCI and canopy N data at lower spatial resolution

In a first step, the relationships between MTCI and canopy N data values were investigated after resampling both datasets to the same lower spatial resolution. The resampled spatial resolutions were 5 km, 10 km, 15 km and 20 km. This was done because of the initial difference in support size between MTCI spatial resolution and the forest plots size (i.e. 1 km and 6 m, respectively). This enabled us to investigate the relationships between MTCI and canopy N data when the spatial discrepancy was accounted for. The statistical basis of this approach is that we

bring both datasets (forest plots and MTCI values) to the same support size or representative area (Bierkens et al. 2000). By averaging out forest plot values within this support size, we calculate the mean of the canopy N value at that support size. By resampling the MTCI values to that same support size, the obtained result consisted of a mean of the MTCI value at that support size. We then regressed the expected canopy N values (at the new support size) against the expected MTCI values (at the new support size).

The monthly averaged MTCI images obtained previously (section 3.2.3.1) were resampled successively to 5 km, 10 km, 15 km, and 20 km. Beforehand, the GlobCover 2009 land cover map was used to exclude from the resampling computation the MTCI pixels located on land surface without natural vegetation cover. As for the forest plots, MTCI pixels whose land cover class corresponded to rain-fed cropland, mosaic between croplands and natural vegetation, sparse vegetation or artificial surfaces were excluded from the upscaling analysis. Forest plot data were then averaged per month over the newly obtained pixel. The relationship between the resampled MTCI values and canopy N data was analysed using log-linear regression.

3.2.3.3 Relationship between MTCI and canopy N data at original spatial resolution (1 km)

In a second step, the relationships between MTCI and canopy N data, both canopy $N_{[\%]}$ and canopy $N_{[\text{area}]}$, were examined at the original spatial resolution of 1 km. This allowed us to investigate the influence of PFTs and species on the relationships as this information was lost in the resampling process. The relationships between MTCI and canopy N at 1 km spatial resolution were analysed with log-linear regression for the whole dataset, for each PFT separately as well as for individual species.

3.2.3.4 Statistical analysis

After applying the selection criteria as explained in the section 3.2.3.1, i.e. plots measured between June 1st and October 31st, exclusion of plots with infrequent species and selection based on GlobCover 2009, 846 forest plots were available for analysis, including 841 plots with foliar biomass and canopy N content information. Descriptive statistics of canopy $N_{[\%]}$, foliar biomass and canopy $N_{[\text{area}]}$ were produced for each of the tree species and PFTs included in the analysis. The log-linear regressions between MTCI and canopy N were performed for both resampled and non-resampled datasets. Preliminary analysis showed that using a natural logarithm transformation (log) of the canopy N variables was necessary to fulfil linear regression model assumptions, namely normality and homogeneity of variance of the residuals. The minimum number of data points needed to carry out the regression analysis was fixed at 10. All the coefficients of determination (r^2) presented are the adjusted r^2 to account for the differences in sample sizes. We calculated the relative root mean squared error of cross-validation (RRMSE_{cv}, %) using the leave-one-out cross-validation method (Clevers and Gitelson, 2013). Its calculation is presented in Equation 3.3 (Yao et al. 2010):

$$RRMSE_{cv} = \sqrt{\frac{1}{n} \times \sum_{i=1}^n (P_i - O_i)^2} \times \frac{100}{\bar{O}_i}, \quad (\text{Equation 3.3})$$

where P_i represents the predicted value, O_i the observed value, \bar{O}_i the mean of all observed value and n the total number of measurements. Resampling both datasets as well as linking the

plots to the MTCI pixels was done with the PCRaster software (Karssenberget al. 2010). The statistical analyses were performed in the R environment (R Development Core Team 2014) and the ggplot2 package was used for the graphics (Wickham 2009).

3.3 Results

3.3.1 Descriptive statistics

Descriptive statistical analysis of canopy $N_{[\%]}$, canopy $N_{[\text{area}]}$ and foliar biomass were performed for each tree species included in the dataset (Table 3.1). The four most abundant species (*Pinus halepensis*, *Pinus sylvestris*, *Quercus ilex* and *Pinus uncinata*) dominated 667 plots, i.e. almost 80% of the plots. The cumulated abundance percentages of ENF, EBF and DBF species were equal to 66 %, 22 % and 9 %, respectively. From these data, it is clear that the forests plots were mainly dominated by ENF species. On average, *Pinus uncinata* plots had the highest biomass values, while *Quercus suber* plots showed the lowest mean value for this variable. Descriptive statistics were also analyzed by PFT. The mean canopy $N_{[\%]}$ was lowest for ENF species (0.97 %N) and highest for DBF trees (2.17 %N) (Figure 3.2a). Canopy $N_{[\%]}$ value ranges were equal to 1.91 %N, 2.06 %N, 1.68 %N and 1.42 %N for DBF, EBF, ENF and mixed plots, respectively. The canopy $N_{[\text{area}]}$ statistics were analysed by PFT as well (Figure 3.2b) and the averaged canopy $N_{[\text{area}]}$ values ranged from 1.82 g m^{-2} to 4.61 g m^{-2} . A Pearson correlation matrix (Figure 3.3) was computed between the variables for the whole dataset. The correlation between each pair of variables was significant and the correlation between canopy $N_{[\text{area}]}$ and foliar biomass was

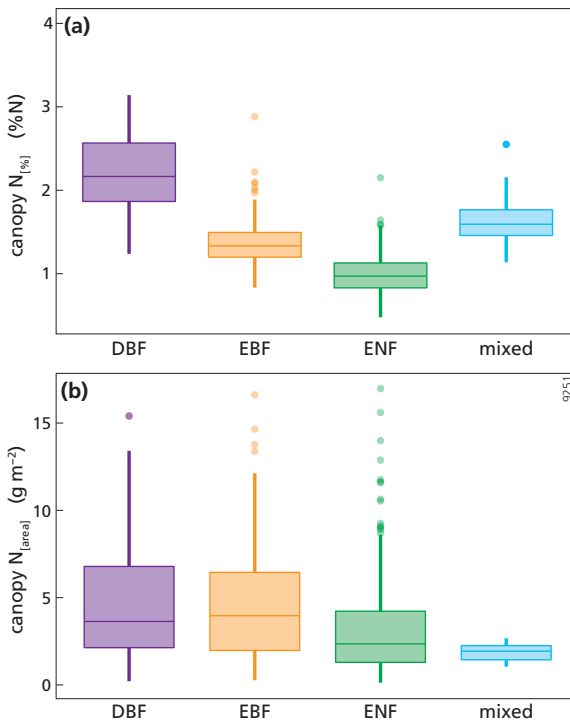


Figure 3.2. Box plots of (a) canopy nitrogen (N) concentration (canopy $N_{[\%]}$, %N) for deciduous broadleaf forest plots (DBF, $n = 80$), evergreen broadleaf forest plots (EBF, $n = 186$), evergreen needleleaf forest plots (ENF, $n = 564$) and mixed forest plots (mixed, $n = 16$); (b) canopy N content (canopy $N_{[\text{area}]}$, g m^{-2}) for deciduous broadleaf forest plots (DBF, $n = 80$), evergreen broadleaf forest plots (EBF, $n = 186$), evergreen needleleaf forest plots (ENF, $n = 563$) and mixed forest plots (mixed, $n = 12$).

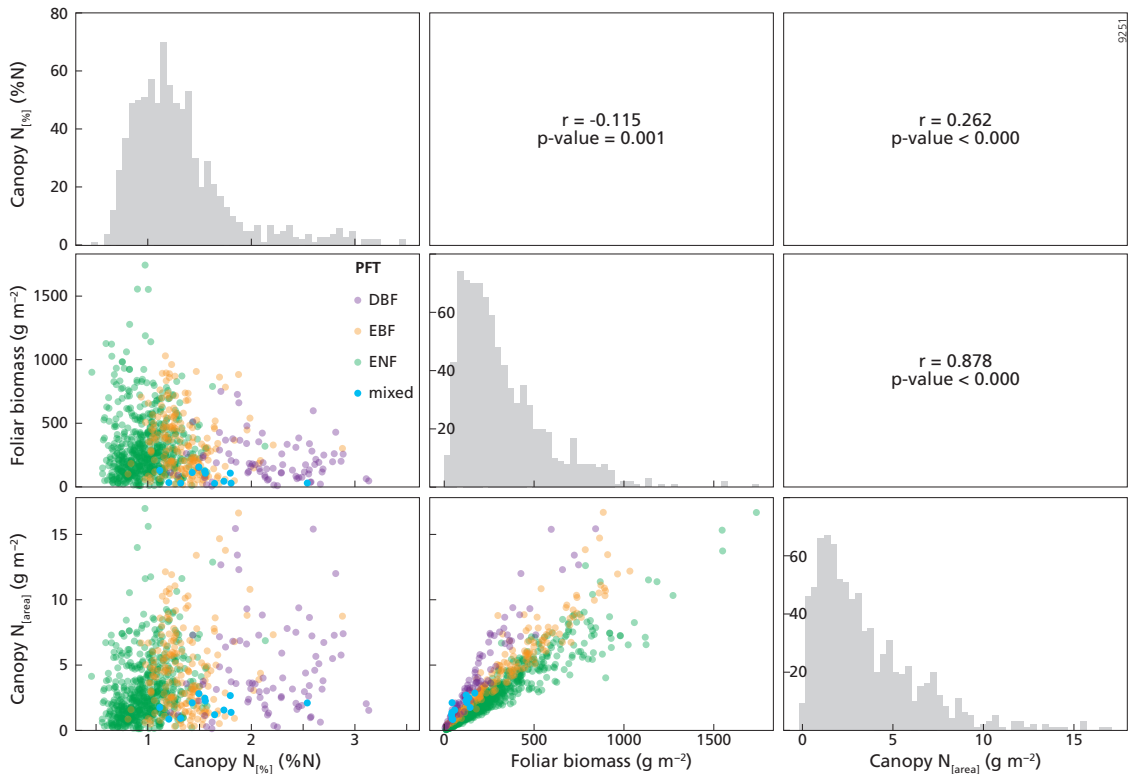


Figure 3.3. The upper right part of this figure shows the Pearson correlation matrix between canopy $N_{[%]}$ (%N), canopy $N_{[area]}$ ($g\ m^{-2}$) and foliar biomass ($g\ m^{-2}$) variables for the whole dataset, $n = 841$. The diagonal presents the histograms of the variables on the x axis, while the y axis represents the number of counts. The lower left part of this figure represents the scatterplots between the variables. PFT is plant functional type, DBF is deciduous broadleaf forest, EBF is evergreen broadleaf forest, ENF is evergreen needleleaf forest, mixed is mixed forest.

strongest ($r = 0.88$). This result was expected as the foliar biomass was included in the $N_{[area]}$ calculation. This matrix also shows distribution histograms of the three variables. As canopy $N_{[%]}$ and canopy $N_{[area]}$ distributions are positively skewed, a logarithmic transformation was applied to these variables to fulfil linear model assumptions. Correlation matrices for each of the DBF, EBF and ENF plots are presented in the Appendix (Figure A.3.2-4).

3.3.2 Relationship between MTCI and canopy N data at lower spatial resolution

The relationships between MTCI and both canopy $N_{[%]}$ and canopy $N_{[area]}$ were studied after resampling both datasets to the same lower spatial resolution. This was done to investigate the relationship between MTCI and canopy N data when the initial spatial discrepancy between the two datasets was accounted for. The results showed that the log-linear relationships between MTCI and either canopy $N_{[%]}$ or canopy $N_{[area]}$ were all highly significant ($p < 0.000$). Moreover, the relationship between MTCI and canopy $N_{[%]}$ was always stronger than the

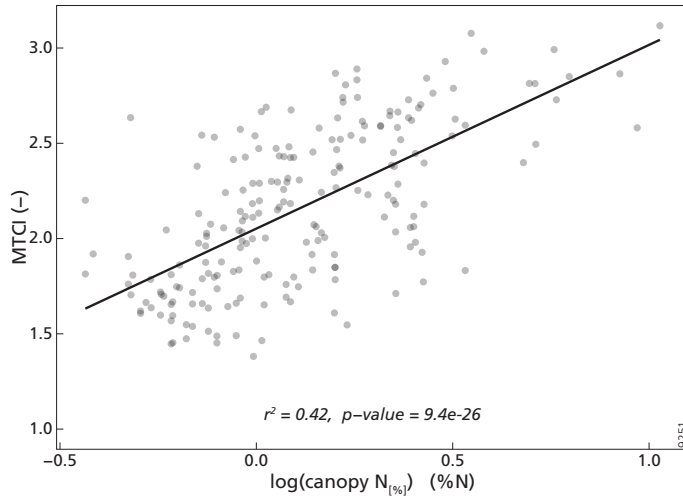


Figure 3.4. Scatterplot between the MTCI (-) and canopy nitrogen concentration (canopy $N_{[%]}$, %N) after resampling the datasets to 20 km spatial resolution ($n = 204$).

relationship for MTCI and canopy $N_{[area]}$ for each resampling factor. The r^2 values of the relationship between MTCI and canopy $N_{[%]}$ were equal to 0.33, 0.37, 0.34 and 0.42 for 5 km, 10 km, 15 km and 20 km resampled spatial resolution, respectively. The r^2 values of the relationship between MTCI and canopy $N_{[area]}$ were equal to 0.20, 0.20, 0.19 and 0.17 at 5 km, 10 km, 15 km and 20 km spatial resolution. The relationship between MTCI and canopy $N_{[%]}$ at 20 km spatial resolution is shown in Figure 3.4. Table 3.2 shows the number of plots per pixel for different pixel sizes (km). As expected, the number of plots per pixel increased with the pixel size, with a mean of 4.1 plots at 20 km spatial resolution. The descriptive statistics of the number of different PFTs, species and sampling years per pixel spatial resolution are provided in the Appendix (Table A. 3.2-4).

Table 3.2. Descriptive statistics of the number of plots per pixel, for different spatial resolutions (km, pixel length): min is minimum, max is maximum, mean is average, sd is standard deviation.

Spatial resolution (km)	Number of plots per pixel			
	min	max	mean	sd
5	1	6	1.44	0.77
10	1	11	2.19	1.53
15	1	15	3.11	2.59
20	1	22	4.09	3.74

3.3.3 Relationship between MTCI and canopy N data at original spatial resolution (1 km)

3.3.3.1 Relationship between MTCI and canopy N concentration

The relationships between MTCI and canopy N data were studied at the original spatial resolution (1 km). The results showed that the log-linear regression between MTCI and canopy $N_{[\%]}$ for the whole dataset ($n = 846$) was highly significant ($p < 0.000$) and had an r^2 value of 0.32 and an RRMSEcv value of 18.7 % (Table 3.3, Figure 3.5a). The relationship between MTCI and canopy $N_{[\%]}$ was also investigated for each PFT individually (Figure 3.5b-e). For DBF plots, the relationship between MTCI and canopy $N_{[\%]}$ had an r^2 value of 0.24 ($n = 80$) and was significant. However, although statistically significant, the r^2 value of the relationship between MTCI and canopy $N_{[\%]}$ for EBF and ENF plots were lower and equal to 0.02 ($n = 186$) and 0.10 ($n = 564$), respectively.

The relationship between MTCI and canopy $N_{[\%]}$ was also significant for one individual species, *Fagus sylvatica*. The proportion of explained variance for this species was equal to 0.69 ($n = 15$). This result, although obtained on a restricted number of plots, showed that the significant relationships between MTCI and canopy $N_{[\%]}$ not only existed when all DBF plots were included but also held for one individual DBF species.

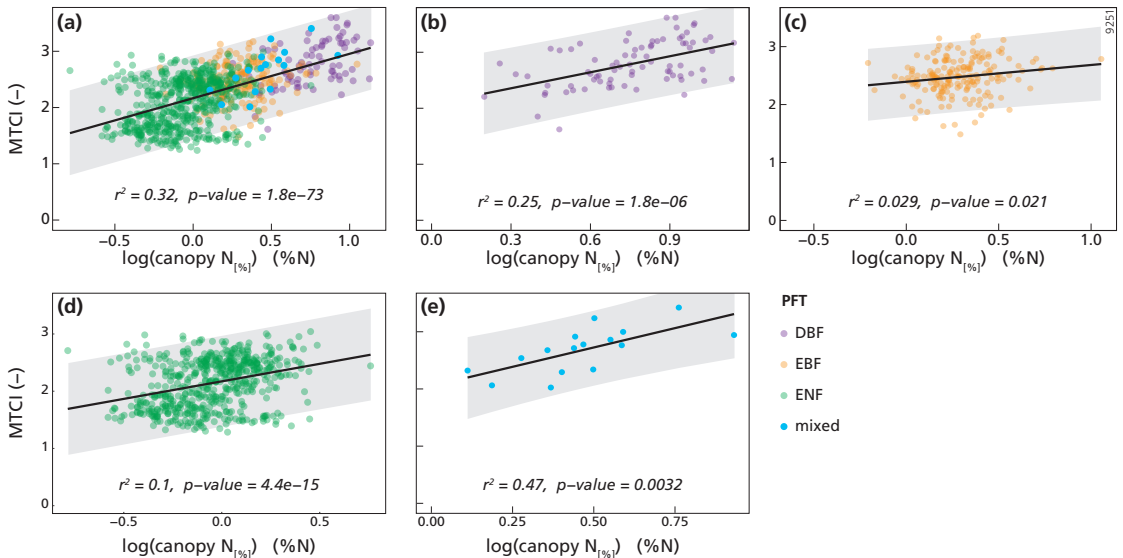


Figure 3.5 Scatterplot and log-linear regression line between the MTCI (-) and canopy nitrogen (N) concentration (canopy $N_{[\%]}$, %N) for (a) the whole dataset ($n = 846$); (b) deciduous broadleaf forest plots (DBF, $n = 80$); (c) evergreen broadleaf forest plots (EBF, $n = 186$); (d) evergreen needleleaf forest plots (ENF, $n = 564$); (e) mixed forest plots ($n = 16$). PFT is plant functional type. The grey shading represents the prediction intervals (95 %). Canopy $N_{[\%]}$ variable was log transformed to fulfil linear model assumptions.

Table 3.3 Observed log-linear regression equations between the MTCI (-) and canopy nitrogen concentration ($CN_{[\%]}$, %N) for different subgroups. Number of plots (n), determination coefficient (r^2), p value and relative root mean squared error of cross-validation (RRMSEcv) are shown. PFT is plant functional type, DBF is deciduous broadleaf forest, EBF is evergreen broadleaf forest, ENF is evergreen needleleaf forest, mixed is mixed forest.

Group	n	Log-linear regression	95% confidence interval intercept	95% confidence interval slope	r^2	p value	RRMSEcv
overall	846	$MTCI = 2.18 + 0.79 \log(CN_{[\%]})$	[2.15, 2.20]	[0.71, 0.87]	0.32	< 0.000	17.0
DBF	80	$MTCI = 2.07 + 0.95 \log(CN_{[\%]})$	[1.78, 2.36]	[0.59, 1.32]	0.24	< 0.000	12.7
EBF	186	$MTCI = 2.39 + 0.29 \log(CN_{[\%]})$	[2.31, 2.48]	[0.04, 0.54]	0.02	0.021	12.4
ENF	564	$MTCI = 2.13 + 0.61 \log(CN_{[\%]})$	[2.10, 2.17]	[0.46, 0.76]	0.10	< 0.000	19.2
mixed	16	$MTCI = 2.05 + 1.35 \log(CN_{[\%]})$	[1.63, 2.46]	[0.53, 2.17]	0.44	0.003	12.4

3.3.3.2 Relationship between MTCI and canopy N content

Significant relationships between MTCI and canopy $N_{[area]}$ were found for the whole dataset as for EBF and ENF plots (Table 3.4). The scatterplots between MTCI and canopy $N_{[area]}$ are presented in Figure 3.6. The proportion of explained variance was higher for ENF plots compared to the other PFTs and compared to the overall relationship across all plots. The

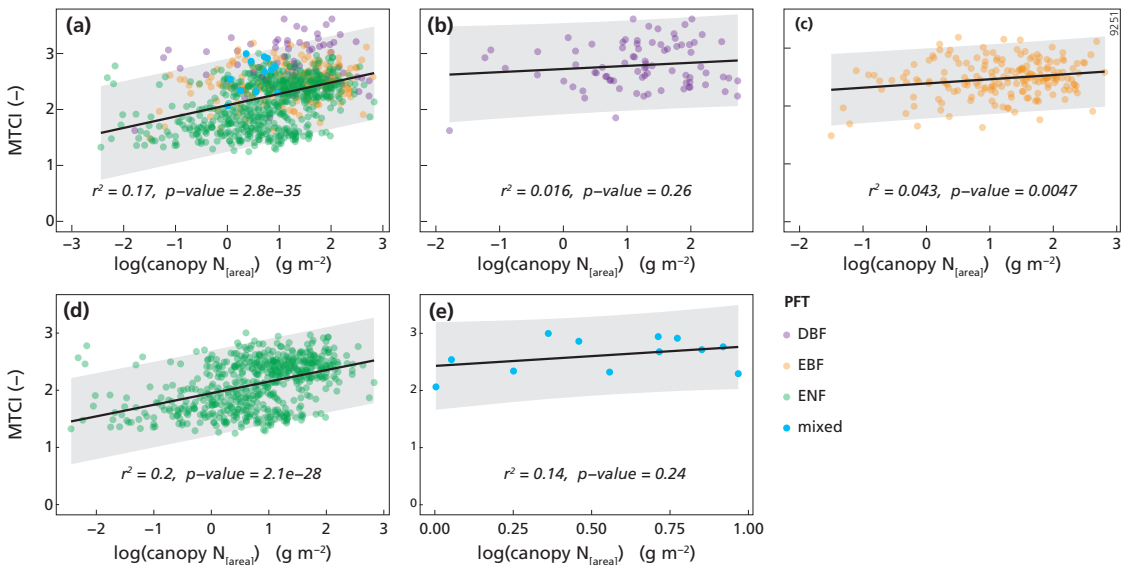


Figure 3.6 Scatterplot and log-linear regression line between the MTCI (-) and canopy N content ($canopy N_{[area]}$, $g m^{-2}$) for (a) the whole dataset ($n = 841$); (b) deciduous broadleaf forest plots (DBF, $n = 80$); (c) evergreen broadleaf forest plots (EBF, $n = 186$); (d) evergreen needleleaf forest plots (ENF, $n = 563$); (e) mixed forest plots ($n = 12$). PFT is plant functional type. The grey shading represents the prediction intervals (95 %). Canopy $N_{[area]}$ variable was log transformed to fulfil linear models assumptions.

Table 3.4 Observed log-linear regressions equations between the MTCI (-) and canopy nitrogen content ($CN_{[area]}$, $g\ m^{-2}$) for different subgroups. Number of plots (n), determination coefficient (r^2), p value and relative root mean squared error of cross-validation (RRMSE_{cv}) are shown. PFT is plant functional type, DBF is deciduous broadleaf forest, EBF is evergreen broadleaf forest, ENF is evergreen needleleaf forest, mixed is mixed forest.

Group	n	Log-linear regression	95% confidence interval intercept	95% confidence interval slope	r^2	p value	RRMSE _{cv}
Overall	841	$MTCI = 2.08 + 0.20 \log(CN_{[area]})$	[2.04, 2.12]	[0.17, 0.23]	0.17	<0.000	18.7
DBF	80	$MTCI = 2.72 + 0.06 \log(CN_{[area]})$	[2.58, 2.87]	[-0.04, 0.15]	0.003	0.263	14.7
EBF	186	$MTCI = 2.39 + 0.07 \log(CN_{[area]})$	[2.32, 2.46]	[0.02, 0.12]	0.04	0.005	12.4
ENF	563	$MTCI = 1.94 + 0.20 \log(CN_{[area]})$	[1.91, 1.99]	[0.17, 0.24]	0.19	<0.000	18.2
mixed	12	$MTCI = 2.43 + 0.34 \log(CN_{[area]})$	[2.05, 2.82]	[-0.26, 0.95]	0.05	0.236	12.8

relationship between MTCI and canopy $N_{[area]}$ was also investigated for 10 individual species and one of them showed significant relationships: *Quercus ilex* ($r^2 = 0.10$, p -value < 0.000, $n = 160$).

3.4 Discussion

3.4.1 Relationship between MTCI and canopy N data at lower spatial resolution

This pre-analysis was undertaken to study the MTCI-canopy N relationships when taking the discrepancy between MTCI original spatial resolution (1 km) and the size of the forest plots (diameter of 6 m) into account. By resampling both datasets to a lower spatial resolution, i.e. 5 km, 10 km, 15 km and 20 km, the obtained values were less impacted by small-scale variations because they were obtained by averaging several values over a larger area. The results showed that the relationship between MTCI and canopy N data was significant and consistent across the resampled spatial resolutions investigated: 5 km, 10 km, 15 km and 20 km. This, however, does not give any indication about the uncertainties resulting from the initial spatial discrepancy between both datasets and about the influence of such uncertainties on the MTCI-canopy N relationship.

3.4.2 Relationship between MTCI and canopy N data at original spatial resolution (1 km)

3.4.2.1 Canopy N concentration

The overall relationship between MTCI and canopy $N_{[\%]}$ at 1 km spatial resolution for all the forest plots ($n = 846$) was significant and the r^2 value was equal to 0.32 (Table 3.3, Figure 3.5). This result showed that canopy $N_{[\%]}$ could be related to MTCI in Mediterranean forests. The performance of the MTCI vegetation index to detect canopy $N_{[\%]}$ in Mediterranean vegetation was similar to the results obtained from previous studies using spaceborne MTCI at higher spatial resolution. For example, by using MTCI computed from the spaceborne RapidEye sensor at 5 m spatial resolution, it was possible to detect canopy $N_{[\%]}$ in a grassland

savannah and subtropical forest with similar coefficients of determination, $r^2 = 0.35$ and $r^2 = 0.52$, respectively (Ramoelo et al. 2012; Cho et al. 2013). However, while there is a consensus regarding MTCI ability for in situ leaf or canopy $N_{[\%]}$ detection in a variety of crops using handheld spectrometers (Tian et al. 2011; Li et al. 2014), there is no general agreement about MTCI ability for canopy $N_{[\%]}$ detection across vegetation and sensor types at larger scales. For example, MTCI computed from airborne data at 3 m spatial resolution could not be related to canopy $N_{[\%]}$ from a mixed temperate forest (Wang et al. 2016). In this context, our finding brings new insight into MTCI $N_{[\%]}$ sensing capabilities at a much coarser spatial resolution (1 km) compared to what has been done before. In these comparisons, it should be taken into account that most previous studies were based on a short sampling campaign while our study incorporates canopy N data from a forest inventory that was carried out during the entire growing season and, therefore, includes differences in phenology.

Investigating the influence of the PFTs on the overall relationship highlighted the difference between DBF, EBF and ENF types of vegetation regarding canopy $N_{[\%]}$ detection by spaceborne MTCI. The relationships between MTCI and canopy $N_{[\%]}$ were significant for all the PFT taken separately (pvalue<0.05). However, a higher proportion of variance was explained for DBF and mixed plots ($r^2 = 0.24$ and $r^2 = 0.44$ for DBF and mixed plots, respectively) compared to the other plant functional types ($r^2 = 0.10$ and $r^2 = 0.02$ for ENF and EBF trees, respectively) and the relationship between MTCI and canopy $N_{[\%]}$ was especially weaker for EBF plots. This indicates that the relationship observed for all the forest plots was mainly driven by DBF and mixed plots. This result is different from what was observed by Ollinger et al. (2008) in boreal forests, where canopy $N_{[\%]}$ was related to NIR reflectance for both broadleaf and needleleaf plots taken separately. Moreover, the results obtained for ENF tree species are surprising as previous studies investigating the relationship between foliar $N_{[\%]}$ and in situ measured spectra reported higher r^2 values, $r^2 = 0.59$ and $r^2 = 0.81$ in spruce and pine forest, respectively (Stein et al. 2014; Schlerf et al. 2010). The differences in scale and methodology might explain the divergent results compared to previous findings. Indeed, in our study, the analysis is carried out at a much coarser spatial resolution using spaceborne data compared to the fine spatial scale obtained with in situ devices. Moreover, most of these studies were carried out in temperate forests and studies investigating canopy $N_{[\%]}$ detection in Mediterranean regions are scarce. When investigating the relationship between canopy $N_{[\%]}$ and MTCI at the species level, we also found that it was significant for *Fagus sylvatica* plots ($r^2 = 0.69$).

In the literature, the relationship between MTCI and canopy $N_{[\%]}$ is often not stratified by PFT or species (Sullivan et al. 2013; Wang et al. 2016). In this study, we showed that investigating this relationship for each PFT taken separately yielded additional insight. Indeed, to our knowledge the difference in explained variance between DBF and other PFTs in the MTCI and canopy $N_{[\%]}$ relationship has not been observed before. Moreover, the results observed for *Fagus sylvatica* plots (n = 15) were consistent with the stronger relationship observed for DBF plots.

3.4.2.2 Canopy N content

The relationship between MTCI and canopy $N_{[\text{area}]}$, which was obtained by combining canopy N concentration values with biomass data, was significant across all plots (n = 841) (Table 3.4, Figure 3.6). Although the r^2 value was lower for the relationship between MTCI and canopy

$N_{[area]}$ ($r^2 = 0.17$) than for the relationship between MTCI and canopy $N_{[%]}$ ($r^2 = 0.32$), it is interesting to note that canopy $N_{[area]}$ can be related to spaceborne MTCI as remotely sensed detection of canopy $N_{[area]}$ is rarely investigated in forest environments (Mirik et al. 2005). In comparison, previous studies conducted in grasslands reported higher prediction accuracy, e.g. by using spaceborne MTCI at 300 m spatial resolution or a simple ratio-type vegetation index computed from airborne imagery at 1 m spatial resolution; canopy $N_{[area]}$ was detected with r^2 values equal to 0.29 and 0.66, respectively (Mirik et al. 2005; Ullah et al. 2012).

The relationship between MTCI and canopy $N_{[area]}$ was only significant for ENF and EBF plots (Figure 3.6b-e), with a higher proportion of explained variance for ENF plots ($r^2 = 0.19$). However, when this relationship was investigated at the species scale, significant results were found for *Quercus ilex* (EBF) plots. This is accordance with a previous study examining the remote sensing of canopy $N_{[area]}$ in *Quercus ilex* trees by MTCI computed from in situ spectra ($r^2 = 0.43$) (Pacheco-Labrador et al. 2014).

3.4.2.3 Comparing results obtained for canopy N concentration and canopy N content

This analysis highlighted the difference between canopy N expressed as a percentage of leaf dry matter (canopy $N_{[%]}$) and on an area basis (canopy $N_{[area]}$) regarding the log-linear relationship with MTCI for the different PFTs. Canopy $N_{[%]}$ of DBF and mixed plots showed higher correlation with MTCI compared to EBF and ENF plots, while the relationship between canopy $N_{[area]}$ of ENF plots with MTCI was stronger than for any other PFTs. These differences between the log-linear relationship between MTCI and either canopy $N_{[%]}$ and canopy $N_{[area]}$ can be related to previous findings showing that canopy $N_{[area]}$ but not canopy $N_{[%]}$ could be detected by MTCI in grassland (Ullah et al. 2012) and by a simple ratio index in heterogeneous rangelands (Mirik et al. 2005) at various spatial scales (300 m and 1 m, respectively). In the literature, canopy $N_{[%]}$ is more often used to detect N state of foliage in forest, while canopy $N_{[area]}$ is regularly employed in grasslands but also in crops (Clevers and Gitelson 2013; Schlemmer et al. 2013). Our results showed that, for ENF plots, when biomass was accounted for, as in canopy $N_{[area]}$, the relationship between MTCI and canopy $N_{[area]}$ was stronger compared to canopy $N_{[%]}$. This suggests that biomass had an influence on and was a confounder of the MTCI-canopy N log-linear relationship.

3.4.3 Possible confounding factors of the MTCI canopy N relationship

The relationships between MTCI and both canopy $N_{[%]}$ and canopy $N_{[area]}$ were influenced by the PFT of the plots. The relationship between MTCI and canopy $N_{[%]}$ was stronger for DBF and mixed plots compared to EBF and ENF plots while the opposite was true for the MTCI-canopy $N_{[area]}$ relationship. In the ongoing discussion about the mechanisms underlying the remote sensing of canopy N, some authors argued that the difference in structural properties between different PFTs was a confounding factor of the observed relationship between canopy N and remote sensing data, rendering it spurious (Knyazikhin et al. 2013). Other authors suggested that the role of canopy structure as confounding factor can be explained by an indirect association between canopy N and canopy structure resulting from convergent adaptive processes (Ollinger et al. 2013; Townsend et al. 2013). In this context, our analysis showed that the PFTs of the plots and the biomass had an influence on the MTCI canopy N relationship in a specific type of ecosystem, namely Mediterranean forests. Other confounding

factors associated with N availability that might affect the observed relationship possibly include biomass, biomass allocation, leaf area index (LAI), water availability and soil type. The data from the forest inventory used in this analysis, i.e. the Catalanian National Forest Inventory, were extensively studied, showing that water availability was the most limiting factor in this region. Water availability was positively correlated with both the $N_{[area]}$ and $N_{[%]}$ in leaves, as well as with foliar and total above-ground biomass through MAP (Sardans et al. 2011; Sardans and Peñuelas 2013). The MAP also influenced the PFT distribution as DBF plots were located in wetter areas than EBF plots, which were found in wetter sites than ENF plots. Regarding the influence of PFTs on the foliar biomass, DBF plots had on average 45% less foliar biomass than EBF or ENF plots (Sardans and Peñuelas 2013).

3.4.4 Perspectives for future applications

The methodology applied in this paper is different from the usual methodology implemented to detect canopy N concentration in forests. Remote sensing of N in forest canopies by hyperspectral sensors is often coupled with intensive forest sampling measurements. This method has been effective for detecting canopy N concentration locally in a vast range of environments (Serrano et al. 2002; Smith et al. 2002; Townsend et al. 2003; Ollinger et al. 2008; Wang et al. 2016). Applying this technique at larger scales has already been explored. For example, Martin *et al.* (2008) compiled 137 field plot data from previous studies in various forest types and investigated the possibility to find a common detection algorithm. However, due to the different treatments required as well as the limited swath width associated with the high spatial resolution (from 3 m to 30 m for HysPex airborne and Hyperion spaceborne sensors, respectively, Wang et al. 2016; Smith et al. 2003), applying imaging spectrometry at a broader scale, although feasible, might reveal to be time consuming. Depending on the sensors as well as on the extent of the study area, this might involve correcting the acquired images for atmospheric influences and cloud cover as well as combining several images into a larger-scale image. A recent study in northern temperate forests explored the effect of spatial resolution on canopy $N_{[%]}$ estimation. The results showed that, although the prediction accuracy was reduced compared to what was achieved using PLS regression at higher spatial resolution, it was still possible to estimate canopy $N_{[%]}$ with r^2 between 0.34 and 0.81 using various vegetation indices computed from Moderate Resolution Imaging Spectroradiometer (MODIS) reflectance data at 500 m spatial resolution (Lepine et al. 2016). In this context, the methodology applied in this article could be a valuable alternative to explore remote sensing of canopy N at larger scale. Using published data from an extensive field plot inventory, we were able to relate both canopy $N_{[%]}$ and canopy $N_{[area]}$ to MTCI at 1 km spatial resolution. Although the relationships found were modest, our study contributes to the ongoing discussion about how to remotely sense canopy N over larger areas. As MTCI time series (1 km) are readily and almost globally available, it could eventually be possible to assess our approach at a broader scale in different types of biomes. The results obtained for DBF species and *Fagus sylvatica* in particular suggest that this method may be efficient for estimating canopy N in temperate forests. If the strength of the relationship between MTCI and canopy N can further be improved, this could lead to canopy N monitoring possibilities at regional scale. In this context, the new OLCI sensor, aboard the Sentinel-3 satellite, and the MultiSpectral Instrument (MSI), aboard Sentinel-2 satellite, might be promising due to their higher spatial resolution (from 10 to 60

m for Sentinel-2). They have bands well positioned to compute the MTCI vegetation index. Although the OTCI, i.e. the successor of the MTCI for the OLCI sensor, is already included in the OLCI level-2b reflectance image, no level 3 product (mosaicked over larger areas and temporally averaged, hence similar to the MTCI time series used in this analysis) is available yet. In addition to more detailed remote sensing data, supplementary ground-based canopy N observations could better constrain the regression models as well. Obtaining reliable ground-based canopy N data over larger areas and for diverse and globally distributed vegetation types would also be necessary to calibrate and validate global vegetation models, as the model performance will depend on the ground data availability and distribution. Remotely sensed canopy N estimates would also improve the calibration of such models. In a recent study, the global vegetation model Lund-Postdam-Jena General Ecosystem Simulator (LPJ-GUESS) was able to simulate the differences in foliar N between different PFTs but not within one PFT (Fleischer et al. 2015). In this context, improving remotely sensed canopy N estimates for homogeneous vegetation types would be a beneficial development for such models.

3.5 Conclusion

In this study, we investigated the relationship between spaceborne MTCI from Envisat and both canopy $N_{[\%]}$ and canopy $N_{[\text{area}]}$ at regional scale in Mediterranean forests. We found significant results across all plots both when the original data were resampled to 5 km, 10 km, 15 km and 20 km and for the original spatial resolution of 1 km. The relationship between MTCI and canopy N data was also significant for some individual PFTs and species. The r^2 values were 0.32 and 0.17 for the overall relationships between MTCI and either canopy $N_{[\%]}$ or canopy $N_{[\text{area}]}$, respectively. We highlighted the differences between PFTs and both canopy $N_{[\%]}$ and canopy $N_{[\text{area}]}$: the relationship between MTCI and canopy $N_{[\%]}$ was stronger for DBF and mixed plots, while canopy $N_{[\text{area}]}$ was more linked to MTCI for ENF plots. Such differences in relationships between MTCI and either canopy $N_{[\%]}$ or canopy $N_{[\text{area}]}$ were already observed in a grasslands ecosystem. Our results showed that MTCI could be related to canopy N for some individual PFTs, indicating an influence of the PFTs on the MTCI-canopy N relationship. The methodology developed in this study could be investigated at larger scales in different types of ecosystems. While this could already be undertaken using the Envisat MTCI 10-year time series as it is almost globally available, ESA's new Sentinel-2 satellite that was launched on 23 June 2015 yields reflectance data at improved spatial and temporal resolution than Envisat-MERIS. Canopy N estimates collected through larger-scales applications could be exploited in vegetation modelling studies including both the C and N cycles.

3.6 Data availability

The canopy data used in this study can be obtained from the TRY Plant Trait Database (<https://www.try-db.org/TryWeb/Home.php>, dataset 91) or by directly contacting the authors.

3.7 Acknowledgements

This research was funded by the Netherlands Organisation for Scientific Research (NWO) under the project number NWO ALW-GO-AO/14-12. J. Sardans and J. Peñuelas were funded by the European Research Council Synergy grant SyG-2013-610028 IMBALANCE-P, the Spanish Government projects CGL2013-48074-P and the Catalan Government project SGR 2014-274. We would like to acknowledge Scott Ollinger and Lucie Lepine for their valuable comments and discussions on our research project as well as Ton Markus for his help with the figures presented in this paper.

3.8 Appendix

This Appendix presents the interannual variation of canopy $N_{[\%]}$ (Figure A. 3.1), the correlation matrices for DBF (Figure A. 3.2), EBF (Figure A. 3.3) and ENF plots (Figure A. 3.4) as well as the tables representing the allometric relationships between foliar biomass and DBH (Table A. 3.1), the number of PFTs (Table A. 3.2), the number of species (Table A. 3.3) and the number of sampling years (Table A. 3.4) per resampled pixel, by pixel spatial resolution.

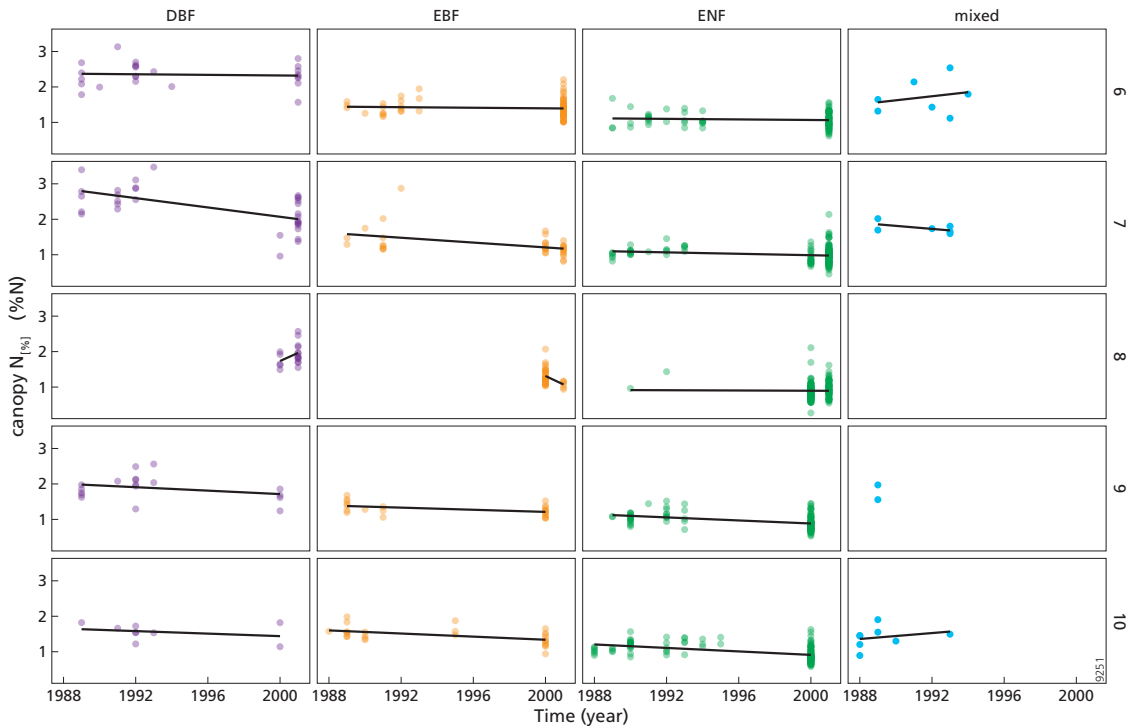


Figure A. 3.1. Interannual variation of canopy $N_{[%]}$ (%N) for each month included in the analysis. The numbers 6-10 (right side of the figure, row numbers) refer to the months of June, July, August, September and October, respectively. DBF is deciduous broadleaf forest, EBF is evergreen broadleaf forest, ENF is evergreen needleleaf forest, mixed is mixed forest. Each point represents an observation at a forest plot. Note that the forest plots were not sampled multiple times; hence, the interannual variation encompasses both temporal variation and spatial variation.

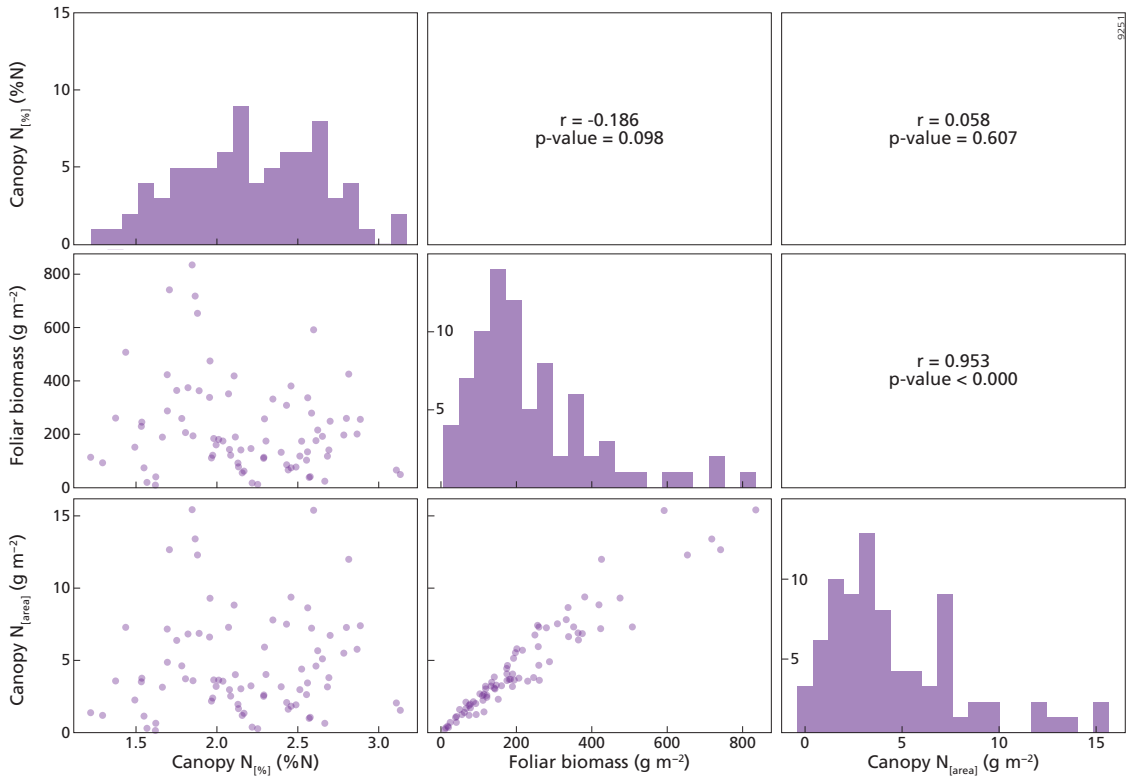


Figure A. 3.2. The upper right part of this figure shows the Pearson correlation matrix between canopy N_[%] (%N), canopy N_[area] (g m⁻²) and foliar biomass (g m⁻²) variables for deciduous broadleaf forest plots (DBF), n = 80. The diagonal presents the histogram of the variable on the x axis, while the y axis represents the number of counts. The lower left part of this figure represents the scatterplots between the variables.

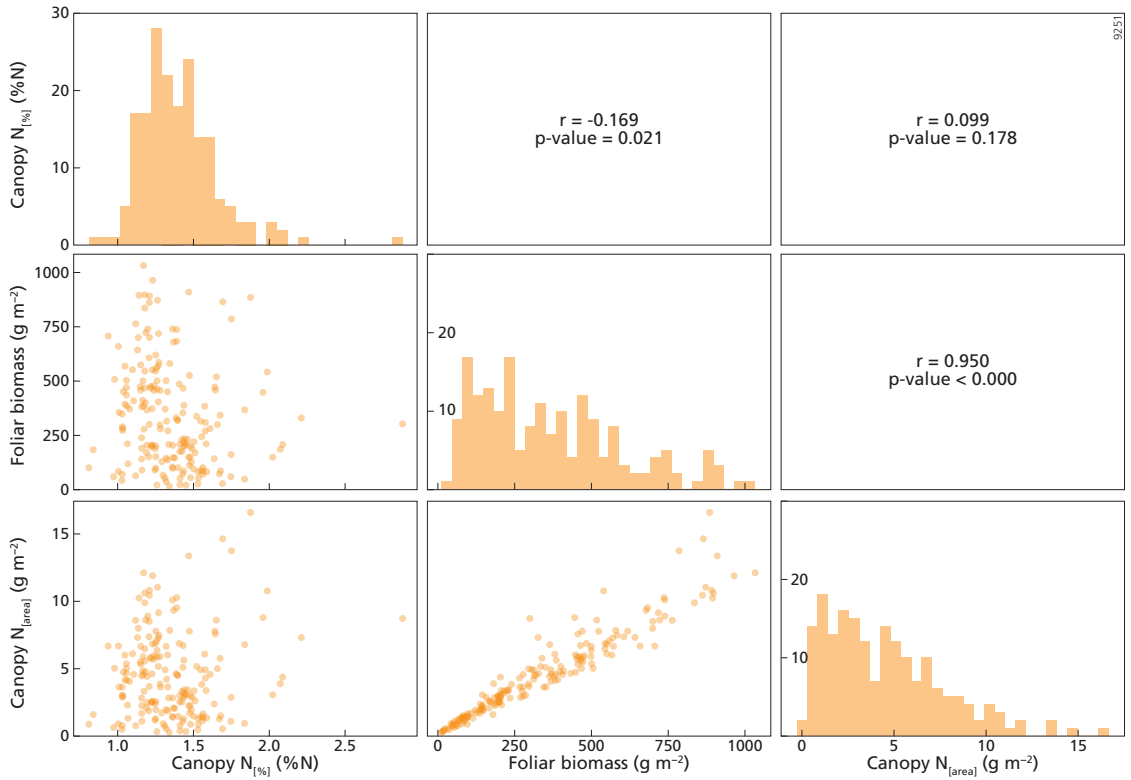


Figure A. 3.3. The upper right part of this figure shows the Pearson correlation matrix between canopy $N_{[\%]}$ (%N), canopy $N_{[area]}$ ($g\ m^{-2}$) and foliar biomass ($g\ m^{-2}$) variables for evergreen broadleaf forest (EBF) plots, $n = 186$. The diagonal presents the histogram of the variable on the x axis, while the y axis represents the number of counts. The lower left part of this figure represents the scatterplots between the variables.

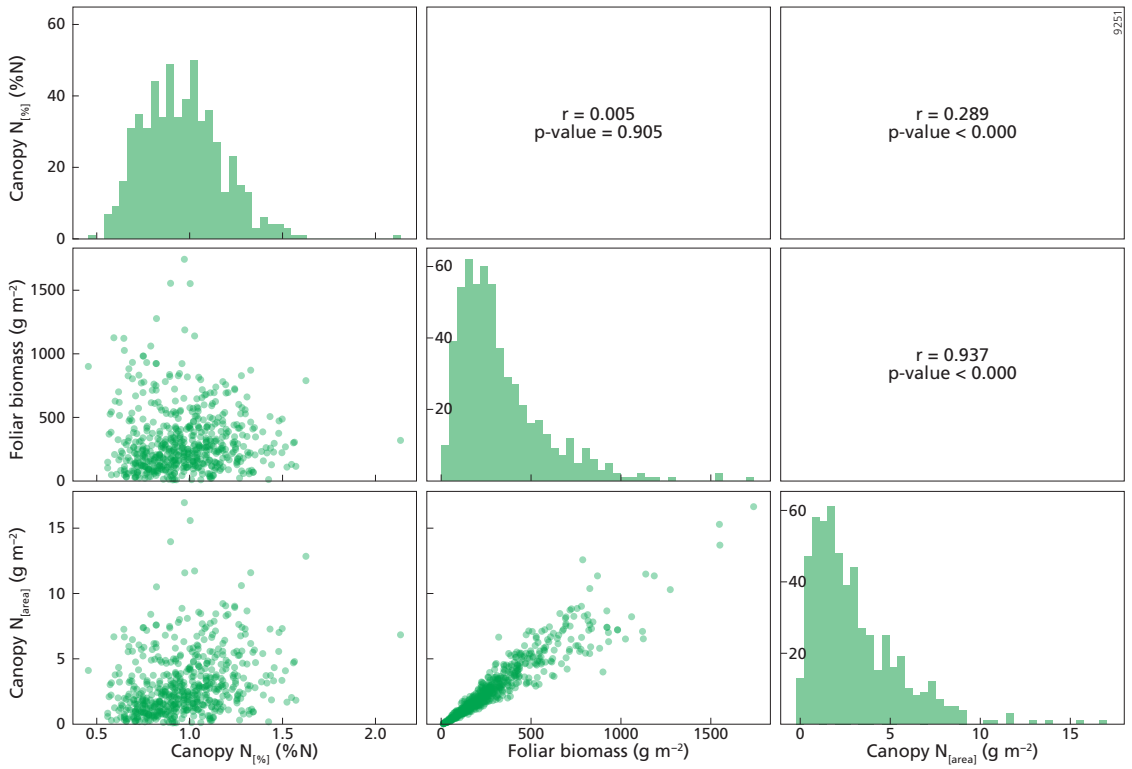


Figure A. 3.4. The upper right part of this figure shows the Pearson correlation matrix between canopy $N_{[%]}$ (%N), canopy $N_{[area]}$ (g m⁻²) and foliar biomass (g m⁻²) variables for evergreen needleleaf forest (ENF) plots, $n = 563$. The diagonal presents the histogram of the variable on the x axis, while the y axis represents the number of counts. The lower left part of this figure represents the scatterplots between the variables.

Table A. 3.1. Allometric relationships between foliar biomass and DBH for the different species included in this analysis. DBH is diameter at breast height (cm). Adapted from (Sardans and Peñuelas, 2015).

Species	Foliar biomass = a · DBH ^b			
	a	b	n	r ²
<i>Castanea sativa</i>	0.032	1.669	86	0.49
<i>Fagus sylvatica</i>	0.026	1.546	285	0.66
<i>Pinus halepensis</i>	0.037	1.656	2420	0.65
<i>Pinus nigra</i>	0.022	1.870	1641	0.65
<i>Pinus pinaster</i>	0.034	1.848	169	0.67
<i>Pinus pinea</i>	0.014	2.029	335	0.72
<i>Pinus sylvestris</i>	0.036	1.651	2755	0.66
<i>Pinus uncinata</i>	0.087	1.410	770	0.62
<i>Quercus canariensis</i>	0.120	1.322	36	0.57
<i>Quercus faginea</i>	0.197	0.943	170	0.40
<i>Quercus humilis</i>	0.047	1.462	595	0.59
<i>Quercus cerroides</i>	0.023	1.805	138	0.73
<i>Quercus ilex</i>	0.063	1.576	2151	0.60
<i>Quercus petraea</i>	0.014	1.888	121	0.73
<i>Quercus suber</i>	0.026	1.446	314	0.55

Table A. 3.2. Descriptive statistics of the number of plant functional types (PFTs) per pixel, by pixel spatial resolution (km): min is minimum, max is maximum, mean is average, sd is standard deviation.

Spatial resolution (km)	Number of PFTs per pixel			
	min	max	mean	sd
5	1	3	1.08	0.29
10	1	4	1.22	0.48
15	1	4	1.34	0.61
20	1	4	1.45	0.69

Table A. 3.3. Descriptive statistics of the number of species per pixel, by pixel spatial resolution (km): min is minimum, max is maximum, mean is average, sd is standard deviation.

Spatial resolution (km)	Number of species per pixel			
	min	max	mean	sd
5	1	4	1.14	0.41
10	1	4	1.38	0.67
15	1	4	1.58	0.85
20	1	6	1.79	1.07

Table A. 3.4. *Descriptive statistics of the number of sampling years per pixel, by pixel spatial resolution (km): min is minimum, max is maximum, mean is average, sd is standard deviation.*

Spatial resolution (km)	Number of sampling years per pixel			
	min	max	mean	sd
5	1	2	1.02	0.15
10	1	3	1.07	0.26
15	1	3	1.10	0.33
20	1	3	1.14	0.40

3.9 References

- Bierkens, M.F.P., Finke, P.A., & De Willigen, P. (2000). *Upscaling and Downscaling Methods for Environmental Research*. Dordrecht: Kluwer Academic Publishers. ISBN: 0-7923-6339-6
- Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., & Ramos Perez, J. (2011). *Globcover 2009 Products Description and Validation Report* (p. 53). Louvain-la-Neuve: UCLouvain & ESA Team. Accessed 2 May 2018. http://due.esrin.esa.int/files/GLOBCOVER2009_Validation_Report_2.2.pdf
- Cho, M.A., Ramoelo, A., Debba, P., Mutanga, O., Mathieu, R., van Deventer, H., & Ndlovu, N. (2013). Assessing the effects of subtropical forest fragmentation on leaf nitrogen distribution using remote sensing data. *Landscape Ecology*, 28, 1479-1491. doi:10.1007/s10980-013-9908-7
- Clevers, J.G.P.W., & Gitelson, A.A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and -3. *International Journal of Applied Earth Observation and Geoinformation*, 23, 344-351. doi:10.1016/j.jag.2012.10.008
- Curran, P.J., & Dash, J. (2005). *Algorithm Theoretical basis document ATBD 2.22 Chlorophyll Index Version 2.2* (p. 41). Southampton: University of Southampton. Accessed 11 November 2016. https://earth.esa.int/documents/700255/2042855/MERIS_ATBD_2.22_v1.2+-+2005.pdf,
- Curran, P.J., Dash, J., Lankester, T., & Hubbard, S. (2007). Global composites of the MERIS Terrestrial Chlorophyll Index. *International Journal of Remote Sensing*, 28, 3757-3758. <https://doi.org/10.1080/01431160600639685>
- Dash, J., & Curran, P.J. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25, 5403-5413. doi:10.1080/0143116042000274015
- Dash, J., & Curran, P.J. (2007). Evaluation of the MERIS terrestrial chlorophyll index (MTCI). *Advances in Space Research*, 39, 100-104. doi:10.1016/j.asr.2006.02.034
- Dash, J., & Vuolo, F. (2010). *Algorithm Theoretical Basis Document OLCI Terrestrial Chlorophyll Index (OTCI) Issue 2* (p. 21). University of Southampton. Accessed 8 March 2018. https://sentinel.esa.int/documents/247904/349589/OLCI_L2_ATBD_OLCI_Terrestrial_Chlorophyll_Index.pdf
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25-36. doi:10.1016/j.rse.2011.11.026
- ESA (2010). *GlobCover 2009 (Global Land Cover Map) Version 2.3* [Dataset]. ESA and UCLouvain. Accessed 11 November 2016. http://due.esrin.esa.int/page_globcover.php
- Evans, J.R. (1989). Photosynthesis and nitrogen relationships in leaves of C3 plants. *Oecologia*, 78, 9-19. doi:10.1007/BF00377192
- Ferwerda, J.G., Skidmore, A.K., & Mutanga, O. (2005). Nitrogen detection with hyperspectral normalized ratio indices across multiple plant species. *International Journal of Remote Sensing*, 26, 4083-4095. doi:10.1080/01431160500181044
- Filella, I., & Penuelas, J. (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *International Journal of Remote Sensing*, 15, 1459-1470. <http://dx.doi.org/10.1080/01431169408954177>
- Filella, I., Serrano, L., Serra, J., & Penuelas, J. (1995). Evaluating wheat nitrogen status with canopy reflectance indices and discriminant analysis. *Crop Science*, 35, 1400-1405. doi:10.2135/cropsci1995.0011183X003500050023x

- Fleischer, K., Wårlind, D., van der Molen, M.K., Rebel, K.T., Arneth, A., Erisman, J.W., Wassen, M.J., Smith, B., Gough, C.M., Margolis, H.A., Cescatti, A., Montagnani, L., Arain, A., & Dolman, A.J. (2015). Low historical nitrogen deposition effect on carbon sequestration in the boreal zone. *Journal of Geophysical Research: Biogeosciences*, 120, 2542-2561. 10.1002/2015JG002988
- Glenn, E.P., Huete, A.R., Nagler, P.L., & Nelson, S.G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*, 8, 2136-2160. doi:10.3390/s8042136
- González, J.R., & Pukkala, T. (2007). Characterization of forest fires in Catalonia (north-east Spain). *European Journal of Forest Research*, 126, 421-429. doi:10.1007/s10342-006-0164-0
- Gracia, C., Ibáñez, J.J., Burriel, J.A., Mata, T., & Vayreda, J. (2004). *Inventari Ecològic i Forestal de Catalunya. Mètodes* (p. 112). Bellaterra: CREAF. Accessed 11 November 2016. <http://www.creaf.uab.es/iefc/pub/Metodes/index.htm>. ISBN: 84-932860-2-8
- Green, D.S., Erickson, J.E., & Kruger, E.L. (2003). Foliar morphology and canopy nitrogen as predictors of light-use efficiency in terrestrial vegetation. *Agricultural and Forest Meteorology*, 115, 163-171. [http://dx.doi.org/10.1016/S0168-1923\(02\)00210-1](http://dx.doi.org/10.1016/S0168-1923(02)00210-1)
- Hansen, P.M., & Schjoerring, J.K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86, 542-553. doi:10.1016/S0034-4257(03)00131-7
- Huber, S., Kneubühler, M., Psomas, A., Itten, K., & Zimmermann, N.E. (2008). Estimating foliar biochemistry from hyperspectral data in mixed forest canopy. *Forest Ecology and Management*, 256, 491-501. doi:10.1016/j.foreco.2008.05.011
- Karssenberg, D., Schmitz, O., Salamon, P., de Jong, K., & Bierkens, M.F.P. (2010). A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. *Environmental Modelling and Software*, 25, 489-502. doi:10.1016/j.envsoft.2009.10.004
- Kergoat, L., Lafont, S., Arneth, A., Le Dantec, V., & Saugier, B. (2008). Nitrogen controls plant canopy light-use efficiency in temperate and boreal ecosystems. *Journal of Geophysical Research: Biogeosciences*, 113. doi:10.1029/2007JG000676
- Knyazikhin, Y., Schull, M.A., Stenberg, P., Möttus, M., Rautiainen, M., Yang, Y., Marshak, A., Latorre Carmona, P., Kaufmann, R.K., Lewis, P., Disney, M.I., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., & Myneni, R.B. (2013). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences*, 110, E185-E192. 10.1073/pnas.1210196109
- Kumar, L., Schmidt, K., Dury, S., & Skidmore, A. (2006). Imaging Spectrometry and Vegetation Science. In F.D.v.d. Meer, & S.M. de Jong (Eds.), *Imaging Spectrometry: Basic Principles and Prospective Applications* (pp. 111-155). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-0-306-47578-8_5. ISBN: 978-0-306-47578-8
- Lana, X., & Burgueño, A. (1998). Spatial and temporal characterization of annual extreme droughts in Catalonia (Northeast Spain). *International Journal of Climatology*, 18, 93-110. 10.1002/(sici)1097-0088(199801)18:1<93::aid-joc219>3.0.co;2-t
- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., & Martin, M.E. (2016). Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sensing of Environment*, 173, 174-186. doi:10.1016/j.rse.2015.11.028
- Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., Liu, Y., Liu, B., Ustin, S.L., & Chen, X. (2014). Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Research*, 157, 111-123. doi:10.1016/j.fcr.2013.12.018

- Liu, D., Ogaya, R., Barbeta, A., Yang, X., & Peñuelas, J. (2015). Contrasting impacts of continuous moderate drought and episodic severe droughts on the aboveground-biomass increment and litterfall of three coexisting Mediterranean woody species. *Global Change Biology*, *21*, 4196-4209. doi:10.1111/gcb.13029
- Martin, M.E., Plourde, L.C., Ollinger, S.V., Smith, M.L., & McNeil, B.E. (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment*, *112*, 3511-3519. doi:10.1016/j.rse.2008.04.008
- 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. doi:10.2111/04-17.1
- NEODC (2015). *NEODC-NERC Earth Observation Data Centre* [Webpage]. UK: Natural Environment Research Council. Accessed 6 February 2015. <http://neodc.nerc.ac.uk/>
- Ollinger, S.V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*, *189*, 375-394. doi:10.1111/j.1469-8137.2010.03536.x
- Ollinger, S.V., Reich, P.B., Frolking, S., Lepine, L.C., Hollinger, D.Y., & Richardson, A.D. (2013). Nitrogen cycling, forest canopy reflectance, and emergent properties of ecosystems. *Proceedings of the National Academy of Sciences*, *110*, E2437. doi:10.1073/pnas.1304176110
- Ollinger, S.V., Richardson, A.D., Martin, M.E., Hollinger, D.Y., Frolking, S.E., Reich, P.B., Plourde, L.C., Katul, G.G., Munger, J.W., Oren, R., Smith, M.L., Paw U, K.T., Bolsta, P.V., Cook, B.D., Day, M.C., Martin, T.A., Monson, R.K., & 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 Sciences of the United States of America*, *105*, 19336-19341. doi:10.1073/pnas.0810021105.
- Ollinger, S.V., & 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*, *8*, 760-778. doi:10.1007/s10021-005-0079-5
- Pacheco-Labrador, J., González-Cascón, R., Pilar Martín, M., & Riaño, D. (2014). Understanding the optical responses of leaf nitrogen in mediterranean holm oak (*Quercus ilex*) using field spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 105-118. doi:10.1016/j.jag.2013.05.013
- Peñuelas, J., Gamon, J.A., Fredeen, A.L., Merino, J., & Field, C.B. (1994). Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sensing of Environment*, *48*, 135-146. doi:10.1016/0034-4257(94)90136-8
- R Development Core Team (2014). *R: A Language and Environment for Statistical Computing* [Computer program]. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>
- Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., & Heitkönig, I.M.A. (2012). Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor. *International Journal of Applied Earth Observation and Geoinformation*, *19*, 151-162. doi:10.1016/j.jag.2012.05.009
- Reich, P.B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 2128-2134. doi:10.1098/rspb.2011.2270
- Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Volin, J.C., & Bowman, W.D. (1999). Generality of leaf trait relationships: A test across six biomes. *Ecology*, *80*, 1955-1969. doi:10.2307/176671

- Reich, P.B., Walters, M.B., & Ellsworth, D.S. (1997). From tropics to tundra: Global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 13730-13734. doi:10.1073/pnas.94.25.13730
- Reich, P.B., Walters, M.B., Kloeppel, B.D., & Ellsworth, D.S. (1995). Different photosynthesis-nitrogen relations in deciduous hardwood and evergreen coniferous tree species. *Oecologia*, *104*, 24-30. doi:10.1007/BF00365558
- Sardans, J., & Peñuelas, J. (2013). Tree growth changes with climate and forest type are associated with relative allocation of nutrients, especially phosphorus, to leaves and wood. *Global Ecology and Biogeography*, *22*, 494-507. doi:10.1111/geb.12015
- Sardans, J., Rivas-Ubach, A., & Peñuelas, J. (2011). Factors affecting nutrient concentration and stoichiometry of forest trees in Catalonia (NE Spain). *Forest Ecology and Management*, *262*, 2024-2034. doi:10.1016/j.foreco.2011.08.019
- Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, *25*, 47-54. doi:10.1016/j.jag.2013.04.003
- Schlerf, M., Atzberger, C., Hill, J., Buddenbaum, H., Werner, W., & Schüller, G. (2010). Retrieval of chlorophyll and nitrogen in Norway spruce (*Picea abies* L. Karst.) using imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, *12*, 17-26. doi:10.1016/j.jag.2009.08.006
- Serrano, L., Peñuelas, J., & Ustin, S.L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*, 355-364. doi:10.1016/S0034-4257(02)00011-1
- Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, *11*, 2027-2054. doi:10.5194/bg-11-2027-2014
- Smith, M.L., & Martin, M.E. (2001). A plot-based method for rapid estimation of forest canopy chemistry. *Canadian Journal of Forest Research*, *31*, 549-555. doi:10.1139/x00-187
- Smith, M.L., Martin, M.E., Plourde, L., & Ollinger, S.V. (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*, *41*, 1332-1337. doi:10.1109/TGRS.2003.813128
- 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. *Ecological Applications*, *12*, 1286-1302. doi:10.2307/3099972
- Stein, B.R., Thomas, V.A., Lorentz, L.J., & Strahm, B.D. (2014). Predicting macronutrient concentrations from loblolly pine leaf reflectance across local and regional scales. *GIScience and Remote Sensing*, *51*, 269-287. doi:10.1080/15481603.2014.912875
- Sullivan, F.B., Ollinger, S.V., Martin, M.E., Ducey, M.J., Lepine, L.C., & Wicklein, H.F. (2013). Foliar nitrogen in relation to plant traits and reflectance properties of New Hampshire forests. *Canadian Journal of Forest Research*, *43*, 18-27. doi:10.1139/cjfr-2012-0324
- Tian, Y.C., Yao, X., Yang, J., Cao, W.X., Hannaway, D.B., & Zhu, Y. (2011). Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance. *Field Crops Research*, *120*, 299-310. doi:10.1016/j.fcr.2010.11.002
- Townsend, P.A., Foster, J.R., Chastain Jr, R.A., & Currie, W.S. (2003). Application of imaging spectroscopy to mapping canopy nitrogen in the forest of the central Appalachian mountains using hyperion

- and AVIRIS. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 1347-1354. doi:10.1109/TGRS.2003.813205
- Townsend, P.A., Serbin, S.P., Kruger, E.L., & Gamon, J.A. (2013). Disentangling the contribution of biological and physical properties of leaves and canopies in imaging spectroscopy data. *Proceedings of the National Academy of Sciences of the United States of America*, 110. 10.1073/pnas.1300952110
- Ullah, S., Si, Y., Schlerf, M., Skidmore, A.K., Shafique, M., & Iqbal, I.A. (2012). Estimation of grassland biomass and nitrogen using MERIS data. *International Journal of Applied Earth Observation and Geoinformation*, 19, 196-204. doi:10.1016/j.jag.2012.05.008
- Vilà-Cabrera, A., Coll, L., Martínez-Vilalta, J., & Retana, J. (2018). Forest management for adaptation to climate change in the Mediterranean basin: A synthesis of evidence. *Forest Ecology and Management*, 407, 16-22. doi:10.1016/j.foreco.2017.10.021
- Vuolo, F., Dash, J., Curran, P.J., Lajas, D., & Kwiatkowska, E. (2012). Methodologies and uncertainties in the use of the terrestrial chlorophyll index for the sentinel-3 mission. *Remote Sensing*, 4, 1112-1133. 10.3390/rs4051112
- Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A.K., Jones, S., Suarez, L., Woodgate, W., Heiden, U., Heurich, M., & Hearne, J. (2016). Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sensing*, 8. doi:10.3390/rs8060491
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag. ISBN: 978-0-387-98140-6
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., & Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, 428, 821-827. doi:10.1038/nature02403
- Yao, X., Zhu, Y., Tian, Y., Feng, W., & Cao, W. (2010). Exploring hyperspectral bands and estimation indices for leaf nitrogen accumulation in wheat. *International Journal of Applied Earth Observation and Geoinformation*, 12, 89-100. <http://dx.doi.org/10.1016/j.jag.2009.11.008>
- Zaehle, S., & Friend, A. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, 24. doi:10.1029/2009GB003521

Chapter 4

Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method



Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method³

Abstract. Canopy nitrogen (N) influences carbon (C) uptake by vegetation through its important role in photosynthetic enzymes. Global Vegetation Models (GVMs) predict C assimilation, but are limited by a lack spatial canopy N input. Mapping canopy N has been done in various ecosystems using remote sensing (RS) products, but has rarely considered environmental variables as additional predictors. Our research objective was to estimate spatial patterns of canopy N in European forests and to investigate the degree to which including environmental variables among the predictors would improve the models compared to using remotely sensed products alone. The environmental variables included were climate, soil properties, altitude, N deposition and land cover, while the remote sensing products were vegetation indices and NIR reflectance from MODIS and MERIS sensors, the MOD13Q1 and MTCI products, respectively. The results showed that canopy N could be estimated both within and among forest types using the random forests technique and calibration data from ICP Forests with good accuracy ($r^2 = 0.62$, RRMSE = 0.18). The predicted spatial pattern shows higher canopy N in mid-western Europe and relatively lower values in both southern and northern Europe. For all subgroups tested (All plots, Evergreen Needleleaf Forest (ENF) plots and Deciduous Broadleaf Forest (DBF) plots), including environmental variables improved the predictions. Including environmental variables was especially important for the DBF plots, as the prediction model based on remotely sensed data products predicted canopy N with the lowest accuracy.

Keywords: Canopy nitrogen, foliar nitrogen, plant traits, ICP Forests, remote sensing, MODIS, MERIS, environmental predictors, random forests, vegetation indices

3 This chapter is based on: Loozen, Y., Rebel, K.T., de Jong, S.M., Lu, M., Ollinger, S.V., Wassen, M.J., & Karssenberg, D. (2020). Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method. *Remote Sensing of Environment*, 247, 111933. <https://doi.org/10.1016/j.rse.2020.111933>

4.1 Introduction

In recent years, mapping canopy nitrogen (N), defined here as the N concentration in plant foliage (g N/100 g dry matter, %N), has been studied at different scales and in a variety of natural environments (Martin et al. 2008; Ollinger et al. 2008; Ramoelo et al. 2012; Wang et al. 2016). This interest in canopy N can be attributed to the role N plays in physiological and ecosystem processes. N is an essential nutrient for plant growth (Chapin 1987). Leaf N concentration is linked to several leaf traits associated with plant photosynthesis (Hikosaka 2004), i.e. photosynthetic capacity (Evans 1989), light use efficiency (Kergoat et al. 2008), specific leaf area and leaf life span (Reich et al. 1999), as shown in the leaf economic spectrum (Wright et al. 2005; Wright et al. 2004) as well as whole-ecosystem net primary productivity (Reich 2012).

Global vegetation models (GVMs) are designed to simulate ecosystem functioning and carbon (C) assimilation by terrestrial ecosystems. Several GVMs explicitly include a representation of the N cycle, which allows them to analyze the influence of the N cycle on the terrestrial C sink (Xu-Ri and Prentice 2008). Spatially explicit data about the N cycle are needed to validate these models. Canopy N mapping through remote sensing could be useful for this purpose.

Mapping canopy N using remote sensing evolved from benchtop studies aiming to identify specific wavelengths related to leaf N concentration using spectroradiometers (Kumar et al. 2006). The red-edge and near infra-red (NIR) have since then been identified as key spectral regions for canopy N estimation (Clevers and Gitelson 2013; Li et al. 2014; Ollinger et al. 2008). The role of the red-edge region, located between 680 and 750 nm (Horler et al. 1983), for canopy N estimation is based on the link between foliar N and chlorophyll through the observed correlation between the red-edge region and leaf chlorophyll content (Clevers and Gitelson 2013; Homolová et al. 2013; Horler et al. 1983; Kokaly et al. 2009; Schlemmer et al. 2013). The NIR spectral region was also identified to correlate with canopy N (Ollinger et al. 2008; Wang et al. 2016). This was observed in temperate and boreal North American forests where canopy N was correlated to both the NIR spectral region as well as NIR-based vegetation indices, including NDVI and EVI (Lepine et al. 2016; Ollinger et al. 2008). Similar relationships were also observed in a mixed European temperate forest (Wang et al. 2016). Although the exact mechanism behind the relationship between canopy N and the NIR reflectance is still unclear, it likely stems from associations between canopy N and the structural properties influencing the NIR scattering.

Among the existing techniques employed for canopy N mapping, creating and using vegetation indices (VIs) is a method that relies on a combination of several spectral regions or bands. Initially developed for crops and local scale applications (Chen et al. 2010; Clevers and Gitelson 2013; Hansen and Schjoerring 2003; Li et al. 2014; Mutanga et al. 2004; Schlemmer et al. 2013; Serrano et al. 2002; Tian et al. 2011), red-edge and, to a lesser extent, NIR-based VIs have been used for canopy N estimation at larger scales in various ecosystems. Several studies focused on grasslands and forest at local scale (Ling et al. 2014; Mirik et al. 2005; Wang et al. 2016) while other studies focused on regional areas such as savannah (Ramoelo et al. 2012) and Mediterranean forests (Loozen et al. 2018).

More recently, environmental variables were used together with remote sensing products to predict canopy N. This approach was suggested by McNeil et al. (2012) who observed an influence of N deposition on the spatial variability of leaf N concentration. Including environmental variables to predict canopy N is thus based on the fact that foliage biochemical concentration is influenced by several environmental factors. In particular, canopy N has been documented to be influenced by climate in Mediterranean forests (Sardans et al. 2011), in Europe (Sardans et al. 2015) and at the global scale (Reich and Oleksyn 2004). Similarly, N deposition affects canopy N (McNeil et al. 2007, 2012; Sardans et al. 2016b; Sardans et al. 2015) as does plant functional type (PFT) (Han et al. 2011; Sardans et al. 2016a; Sardans et al. 2015). Soil properties, i.e. soil pH and nutrients, were also found to correlate with canopy N (Han et al. 2011). This approach, i.e. including environmental variables to predict canopy N using remote sensing, was implemented in a study mapping canopy N in savannah grass using red-edge VIs as well as several environmental variables: soil, climate, geology and altitude (Ramoelo et al. 2012). In a recent study, Moreno-Martínez et al. (2018) used the random forests algorithm to map canopy N at global scale for several PFTs. As predictor variables, they used bands and VI products from the MODIS sensor as well as environmental variables, i.e. bioclimatic variables, surface temperature and elevation.

The random forests algorithm used by Moreno-Martínez et al. (2018), is a machine learning technique based on regression trees which allows to model nonlinear relationships using several types of explanatory variables. It was found to be among the best techniques to predict foliar traits (Moreno-Martínez et al. 2018). Random forests have mainly been implemented in grasslands, at local (Adjorlolo et al. 2014; Mutanga et al. 2015) and regional scales (Ramoelo et al. 2015), but also in a coffee plantation (Chemura et al. 2018) and in the miombo woodlands (Mutowo et al. 2018). These studies included either all reflectance bands available or several VIs as predictor variables for canopy N.

In this context, although several studies attempted to develop a methodology to map canopy N over large spatial extents (Lepine et al. 2016; Martin et al. 2008; Moreno-Martínez et al. 2018), no study so far investigated the feasibility of mapping the spatial patterns of canopy N in European forests. In this study, our research objective was (i) to predict canopy N and its spatial pattern over European forests and (ii) to test whether including environmental variables as predictors improves canopy N predictions compared to approaches that rely on remotely sensed data alone. To do so, we mapped canopy N in European forests using the canopy N data from the ICP network as calibration data. We related canopy N plot data to the NDVI, EVI and NIR obtained from the MODIS MOD13Q1 product and the MTCI from the MERIS sensor, and environmental variables. The environmental variables included were elevation, climate, soil properties, N deposition and land-cover. We used the random forests machine learning technique to relate canopy N to the predictor variables. To evaluate the influence of including environmental variables on the results, we evaluated nine different random forests model settings: models using all predictor variables (*All pred*), using only remote sensing variables (*RS only*), and using only environmental variables (*Env only*). Each model was parameterized on three subgroups: all available plots (All plots), only Evergreen Needleleaf Forest (ENF) plots and only Deciduous Broadleaf Forest (DBF) plots. Including these three subgroups provided insights about the feasibility of mapping canopy N at European scale on all available plots, ENF plots and DBF plots. The results of the models were evaluated on each subgroup separately.

4.2 Material and methods

4.2.1 Canopy N data

4.2.1.1 ICP Forests

Canopy N data used in this analysis were obtained from the ICP Forests program (International Co-operative Program on Assessment and Monitoring of Air Pollution Effects on Forests, www.icp-forests.net). ICP Forests is a European biomonitoring network of forest conditions. The intensive monitoring program (level II network) includes more than 800 permanent forest plots sampled regularly across European countries. The forest plots are in homogeneous forest sites selected such that the diversity in European forests is represented. The forest plots have a minimum size of 0.25 ha, which corresponds to 56 m diameter for a circular plot (Ferreti et al. 2017). The foliar chemistry survey, including canopy N measurements, has been repeated every two years. The plots were sampled following a standard and consistent sampling design. Minimum five trees of each species belonging to the dominant class were selected. The sampling was repeated on the same sampled trees over the years. The leaves or needles were collected from the upper third part of the crown. If several species composed the dominant forest class, the foliar chemistry analysis was done separately for each species. Deciduous species plots were sampled during the second half of the growing season, before the onset of autumn, while evergreen plots were sampled during winter months, in the dormancy period. Quality control of the foliar concentration measurement was ensured by means of regular interlaboratory comparisons (Rautio et al. 2016).

4.2.1.2 Canopy N data analysis

Annual plot canopy N measurements data were obtained from the ICP Forests website for the period 1990-2014. Missing and duplicate entries as well as rare tree species, i.e. species that were sampled in less than six plot measurements, were excluded from the analysis. Canopy N outlier values, defined as those that were outside of the species-specific 5-95% percentile, were also removed from the dataset. 5207 annual plot measurements were left for analysis. Canopy N annual plot measurements were averaged by plot over all the sampling years to produce long-term averages of plot canopy N. This represented 818 plots, for which we obtained a long-term average canopy N value. Long-term averages canopy N will be called canopy N in the rest of the article. Plots were labelled according to their PFT. Plots with trees belonging to different PFTs were labelled as mixed PFT. Descriptive statistics of the canopy N data were performed.

4.2.2 Environmental variables

We chose to include as predictor variables the following environmental variables for their known influences on the N cycle and ecosystem properties in general.

4.2.2.1 Bioclimatic variables

Climate was found to be related to canopy N (Reich and Oleksyn 2004; Sardans et al. 2015; Sardans et al. 2011). Bioclimatic variables from the WorldClim2 dataset (Fick and Hijmans 2017) were used. The bioclimatic variables were computed from monthly temperature and precipitation over the period 1972-2000. The bioclimatic variables consist in annual mean,

seasonality, minimum or maximum values. The complete list of bioclimatic variables (19) included in the analysis is presented in Table 4.1. The initial spatial resolution was 1 km.

4.2.2.2 Altitude

We have decided to include elevation as a predictor variable because of both the correlation found between canopy N and temperature at global scale (Reich and Oleksyn 2004) and the relationship between altitude and temperature. The digital elevation model over Europe (EU-DEM) was used for altitude data (European Environment Agency 2013). The EU-DEM is a digital surface model based on both SRTM and ASTER GDEM as source data. The EU-DEM was produced using Copernicus data. The EU-DEM was obtained from the European Environmental agency website for the extent of Europe. The original spatial resolution was 30 m.

4.2.2.3 Soil properties

We chose to include soil properties as predictor variables because soil is an important component of the ecosystem that influences vegetation. More specifically, canopy N has been shown to be correlated to soil pH and soil mineral content (Han et al. 2011). Soil property maps were obtained from Soilgrids250m Global Soil Information (Hengl et al. 2017). The Soilgrids250m soil properties maps were predicted from a large collection of soil profile samples and globally available remote sensing products using machine learning prediction techniques. The list of Soilgrids250m variables included in this analysis (e.g. soil pH, soil cation exchange capacity, CEC, soil sand, clay and silt content, soil depth) is presented in Table 4.1. The original spatial resolution was 250 m.

4.2.2.4 N deposition

Canopy N has been found to be correlated with N deposition in various ecosystems including in *Pinus sylvestris* forests and mixed European forests (McNeil et al. 2007, 2012; Sardans et al. 2016b; Sardans et al. 2015). N deposition was included as predictor variable. The N deposition maps used in this analysis were aggregated from three atmospheric chemistry models (GISS-E2-R, CCSM-CAM3.5 and GFDL-AM3, Lamarque et al. (2013a)) within the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP, Lamarque et al. (2013b)). The maps were obtained from the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) website (ISIMIP 2019). We used reduced (NH_x), oxidized (NO_y) and total N deposition maps for the year 2006. The total N deposition was calculated as the sum of oxidized and reduced N depositions. The initial spatial resolution was 0.5 x 0.5 decimal degrees or approximately 50 km.

4.2.2.5 Land cover

We included the land cover as predictor variable in this analysis because the plant functional type has been shown to be related to leaf N (Kattge et al. 2011; Sardans et al. 2016a; Sardans et al. 2015). We used the land cover (LC) map from the ESA climate change initiative (CCI) over the epoch 2008-2012 (v.1.6.1, Defourny et al. (2016)). Following the UN land cover classification system, the land cover includes 22 land cover classes, which are compatible with the classification used in GVMs. We chose the ESA CCI LC from the epoch 2008-2012 because

it is the period with the highest number of annual plot canopy N measurements. The LC spatial resolution was 300 m.

4.2.3 Remote sensing variables

4.2.3.1 MOD13Q1 product

MODIS data included in this analysis were NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index) VIs as well as NIR reflectance from the MOD13Q1 product (Didan 2015). We chose to include these remote sensing products because the relationship between canopy N and NIR, either as a stand-alone reflectance product or included in the calculation of the NDVI and EVI, is well documented (Chemura et al. 2018; Mutowo et al. 2018; Ollinger et al. 2008). The MOD13Q1 product is available globally for every 16 days period at 250 m spatial resolution. One MODIS image for each 16 days period between the 1st January 2002 and the 31st December 2014 was obtained for each product considered (NIR, NDVI and EVI) as well as for the pixel reliability quality layer (QA). The MODIS images were downloaded from the AppEARS website (AppEARS Team 2019) for the extent of Europe.

4.2.3.2 MTCI product

The MTCI (MERIS Terrestrial Chlorophyll Index) was originally developed to monitor chlorophyll content in vegetation (Dash and Curran 2004). It has been related to canopy N in various types of ecosystem and for various spatial extents, from local to regional studies (Cho et al. 2013; Loozen et al. 2018; Ramoelo et al. 2012; Tian et al. 2011). MTCI is a red-edge based VI that is computed using three MERIS sensor's reflectance bands located near the red-edge region (Equation 4.1).

$$MTCI = \frac{R_{\text{band10}} - R_{\text{band9}}}{R_{\text{band9}} - R_{\text{band8}}} = \frac{R_{753.75} - R_{708.75}}{R_{708.75} - R_{681.75}}, \quad (\text{Equation 4.1})$$

The MTCI level 3 product is available almost globally as a monthly average at 1 km spatial resolution. The original reflectance data were provided by the European Space agency and were processed by Airbus Defense and space. The MTCI imagery is distributed by the Natural Environment Research Council (NERC) Earth Observation Data Centre (NEODC 2015). One MTCI image was downloaded for the extent of Europe for each month between June 2002 and December 2011, except for October 2003, as no valid product was available.

4.2.4 Data preprocessing

Both MODIS and MTCI imagery were averaged by month to produce 12 long-term monthly averaged images for each remote sensing product considered. Before averaging, pixel-based quality information from the QA layers were applied on MODIS images. All pixels for which the quality value was not labelled as 'good' were excluded from the computation of the long-term average. While there is no quality layer available for MTCI imagery, MTCI values lower than 1 are not valid. We ensured that no pixel with invalid values were included in the long-term MTCI monthly averages. However, by doing so, virtually no MTCI pixel were excluded from the calculation. The obtained long-term monthly averages of the MODIS products

in January, February and December as well as the MTCI product in January and December contained a high number of pixels with missing values. These long-term monthly averages were excluded from the analysis.

Among the 22 LC classes initially present in the CCI LC variable, some were rarely represented in our study area. We grouped together similar LC classes and obtained 15 LC classes. Each of the obtained LC class was converted to a binary layer, in which the pixel values corresponded to the presence or absence of the specified LC class. All of the binary layers obtained were used as variables in the analysis.

All the predictor variables layers, including MODIS and MTCI long-term monthly averages, were resampled to a common grid and spatial resolution of 300 m using the bilinear interpolation of the resample function of the raster package in the R environment (Hijmans 2018; R Core Team 2019).

4.2.5 Random forests

Random forests is a machine learning method that is built on the classification and regression trees (CART) with the ensemble method. It was developed by Breiman (2001) and has been applied to map canopy N using VIs and other remote sensing products at different scales and in different ecosystems (Chemura et al. 2018; Moreno-Martínez et al. 2018; Mutowo et al. 2018; Ramoelo et al. 2015).

Random forests avoids overfitting by randomly sampling the predictor space. It is can model non-linear relationships without being constrained by the assumptions of the variable distributions and dependency. In a recent study, random forests was also found to give better leaf trait predictions compared to regularized linear regression, neural networks and kernel methods (Moreno-Martínez et al. 2018).

Random forests works by training many regression trees and reporting the mean response over all the trees. We implemented the random forests analysis in the R environment (R Core Team 2019) using the randomForest package (Liaw and Wiener 2002). The random forests algorithm is governed by three parameters, the number of trees (ntree), the number of sampled variables (mtry) and the minimum number of terminal seeds (nodesize). The random forests algorithm and the settings of the models are as follows:

- The regression tree is grown by iteratively splitting the bootstrap sample into two groups using the best predictor from a randomly selected subsample of all the available predictors. The mtry was set to one third of the total number of predictor variables.
- To build each regression tree, a bootstrap sample including two third of the training data is randomly selected. The remaining third of the training data (called the out-of-bag data (OOB)) is used to evaluate this specific tree.
- The tree is grown until the nodesize is reached. We set the node size to 5.
- This process repeated for ntree number of times. The ntree parameter was set to 2500 trees.

The random forests algorithm provides the mean squared errors and r-squared values assessed using the OOB samples, MSE_{OOB} and r^2_{OOB} , respectively.

We implemented the random forests model to predict canopy N in European forests. We fitted the random forests model to long-term average plot canopy N (section 4.2.1.2) using

Table 4.1. List of all the predictors variables, both environmental and remote sensing variables, included in the analysis. For the remote sensing variables MOD13Q1 and MTCI, there is a long-term monthly average variable for each month. There are not listed here for the sake of length.

Type of variable	Variable set	Variable name	Unit	Original spatial resolution	Source		
Environmental	Bioclimatic variable	Annual mean temperature	°C	1 km	Fick and Hijmans, 2017		
		Mean diurnal range (mean of monthly (max temp - min temp))	°C	1 km	Fick and Hijmans, 2017		
		Temperature seasonality (standard deviation *100)	°C	1 km	Fick and Hijmans, 2017		
		Max temperature of warmest month	°C	1 km	Fick and Hijmans, 2017		
		Min temperature of coldest month	°C	1 km	Fick and Hijmans, 2017		
		Temperature annual range	°C	1 km	Fick and Hijmans, 2017		
		Isothermality (mean diurnal range/temp annual range)	%	1 km	Fick and Hijmans, 2017		
		Mean temperature of wettest quarter	°C	1 km	Fick and Hijmans, 2017		
		Mean temperature of driest quarter	°C	1 km	Fick and Hijmans, 2017		
		Mean temperature of warmest quarter	°C	1 km	Fick and Hijmans, 2017		
		mean temperature of coldest quarter	°C	1 km	Fick and Hijmans, 2017		
		Annual precipitation	mm	1 km	Fick and Hijmans, 2017		
		Precipitation of wettest month	mm	1 km	Fick and Hijmans, 2017		
		Precipitation of driest month	mm	1 km	Fick and Hijmans, 2017		
		Precipitation seasonality (coefficient of variation)	%	1 km	Fick and Hijmans, 2017		
		Precipitation of wettest quarter	mm	1 km	Fick and Hijmans, 2017		
		Precipitation driest quarter	mm	1 km	Fick and Hijmans, 2017		
		Precipitation warmest quarter	mm	1 km	Fick and Hijmans, 2017		
		Precipitation of coldest quarter	mm	1 km	Fick and Hijmans, 2017		
		Soil properties	Altitude	EU DEM altitude	m	30 m	European Environment Agency, 2013
				Absolute depth to bedrock (BDTCM)	cm	250 m	Hengl et al. 2017
				Bulk density (fine earth, BLDfIE)	kg m ⁻³	250 m	Hengl et al. 2017
				Soil pH in water (PHIHOX)	pH	250 m	Hengl et al. 2017
Cation exchange capacity of soil (CECSOL)	cmolc kg ⁻¹			250 m	Hengl et al. 2017		
Coarse fragments volumetric (CRFVOL)	%			250 m	Hengl et al. 2017		
Sand content (50–2000 micro meter) mass fraction (SNDPPT)	%			250 m	Hengl et al. 2017		
Silt content (2–50 micro meter) mass fraction (SLTPPT)	%			250 m	Hengl et al. 2017		
Clay content (0–2 micro meter) mass fraction (CLYPPT)	%			250 m	Hengl et al. 2017		
Soil organic carbon stock (OCSTHA)	tons ha ⁻¹			250 m	Hengl et al. 2017		
Nitrogen deposition	CCI Land cover	Soil organic carbon density (OCDENS)	kg m ⁻³	250 m	Hengl et al. 2017		
		Soil organic carbon content (fine earth fraction)	g kg ⁻¹	250 m	Hengl et al. 2017		
		Oxidized nitrogen deposition (NOy)	g N m ⁻² yr ⁻¹	50 km	Lamarque et al. 2013a		
		Reduced nitrogen deposition (NHx)	g N m ⁻² yr ⁻¹	50 km	Lamarque et al. 2013a		
		Total nitrogen deposition	g N m ⁻² yr ⁻¹	50 km	Lamarque et al. 2013a		
		CCI land cover map (epoch 2010)	factorial	300 m	Defourny et al. 2016		
Remote sensing	MOD13Q1	MODIS NDVI long-term monthly averages	/	250 m	Didan, 2015		
		MODIS EVI long-term monthly averages	/	250 m	Didan, 2015		
		MODIS NIR long-term monthly averages	/	250 m	Didan, 2015		
		MTCI long-term monthly averages	/	1 km	Dash and Curran, 2004		

the predictor variables. We tested nine different models. A first type of model included all the predictor variables (*All pred*), a second type of model included only the remote sensing variables (*RS only*) and a third type of model included only the environmental variables (*Env only*). Each type of model was tested separately on three subgroups: All plots, only ENF plots and only DBF plots. We included these three subgroups in this study because each group provides different insight about canopy N spatial patterns and the feasibility of mapping canopy N at European scale. We did not develop a separate model for either EBF or mixed plots because of the restricted number of plots for these two PFTs (29 and 11, respectively).

The subsequent workflow was applied for all the models tested in this analysis:

- A first random forests model was fitted to the canopy N data. The predictor variables included in the model were selected using a recursive backward elimination (Brungard et al. 2015; Mutanga et al. 2015). The model was first fitted with all the predictors. The least important predictor was removed from the model. This process was repeated until only one predictor variable was left. The model selected was the one with lowest MSE_{OOB} value. r^2_{OOB} and $RRMSE_{OOB}$ are reported.
- We used 10-fold cross-validation (C-Val) as independent validation to assess the accuracy of the selected model. Theand are reported. The C-Val is calculated using the R caret package (Kuhn 2018).

The coefficient of determination (r^2) is calculated as

$$1 - \frac{MSE}{Var(CN)}, \quad (\text{Equation 4.2})$$

where MSE is the mean squared error, Var is the variance and CN, the canopy nitrogen (%N). The Relative Root Mean Squared Error (RRMSE) is calculated as

$$\sqrt{\frac{1}{n} \times \sum_{i=1}^n (P_i - O_i)^2} \times \frac{1}{\bar{O}_i}, \quad (\text{Equation 4.3})$$

where $i = 1, 2, \dots, n$ are distinct values, n is the total number of values, P_i is the predicted value, O_i is the observed value and \bar{O}_i is the mean of all observed values.

The importance of the predictor variables is assessed by randomly permuting each predictor variable and calculating the subsequent decrease in OOB accuracy. The importance measure is expressed as the mean decrease in MSE_{OOB} (Liaw and Wiener 2002).

4.2.6 Mapping canopy N

The best model for each subgroup (All plots, DBF and ENF), assessed using r^2_{CV} and $RMSE_{CV}$, is used as predictive model to map canopy N in European forests.

4.3 Results

4.3.1 Descriptive analysis of canopy N plot data

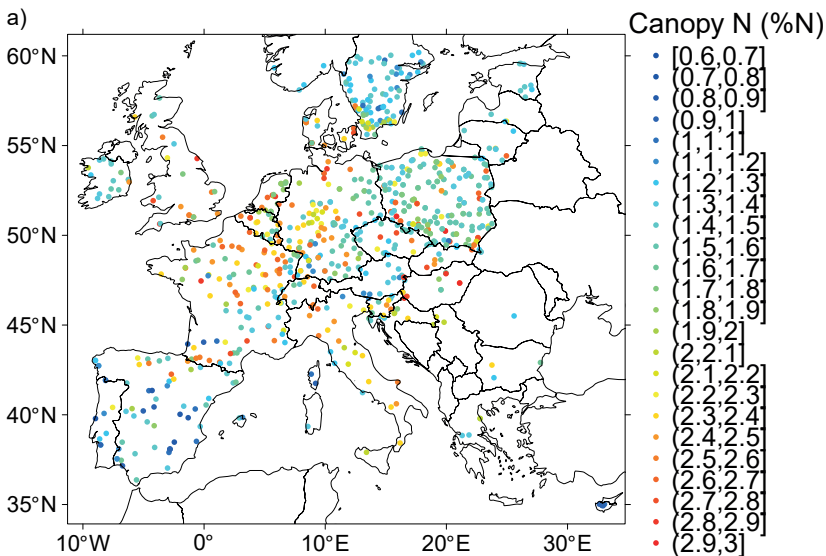
Table 4.2 gives the canopy N plot data descriptive statistics. Among the 818 forest plots included in the analysis, the majority belongs to the ENF forest type (63 %), while the second

most occurring PFT, i.e. DBF, is present in 32 % of the plots. Forest plots where several PFTs were sampled represent only 1% of the total. As expected, ENF plots have on average lower canopy N compared to DBF plots (1.4 and 2.4 %N, respectively). The locations of the canopy N plot data are presented in Figure 4.1. Higher canopy N values are observed in Midwestern Europe while lower canopy N values occur in the southern and northern part of the study region (Spain and south of Sweden, respectively, Figure 4.1a). Regarding the PFTs, DBF and ENF occur in the whole study area, while EBF and mixed plots are only found in certain regions, i.e. in Southern Europe and in central Western Europe, respectively. The observed difference in canopy N values between ENF and DBF plots is reflected in the respective PFTs maps (Figure 4.1 – b and c).

Table 4.2. Descriptive statistics of long-term average canopy nitrogen concentration (%N) plot data calculated over all forest plots (All plots) and grouped by plant functional type (PFT): Deciduous Broadleaf Forest (DBF), Evergreen Broadleaf Forest (EBF), Evergreen Needleleaf Forest (ENF) and mixed forest plots (mixed), with minimum (min), maximum (max), mean, and standard deviation (sd) values.

PFT	Number of plots (%)	Canopy N (%N)			
		min	max	mean	sd
All plots	818 (100 %)	0.6	3.0	1.8	0.5
DBF	265 (32 %)	1.5	3.0	2.4	0.2
EBF	29 (4 %)	1.2	1.6	1.4	0.1
ENF	513 (63 %)	0.6	2.2	1.4	0.2
mixed	11 (1 %)	1.4	2.4	1.7	0.3

Figure 4.1



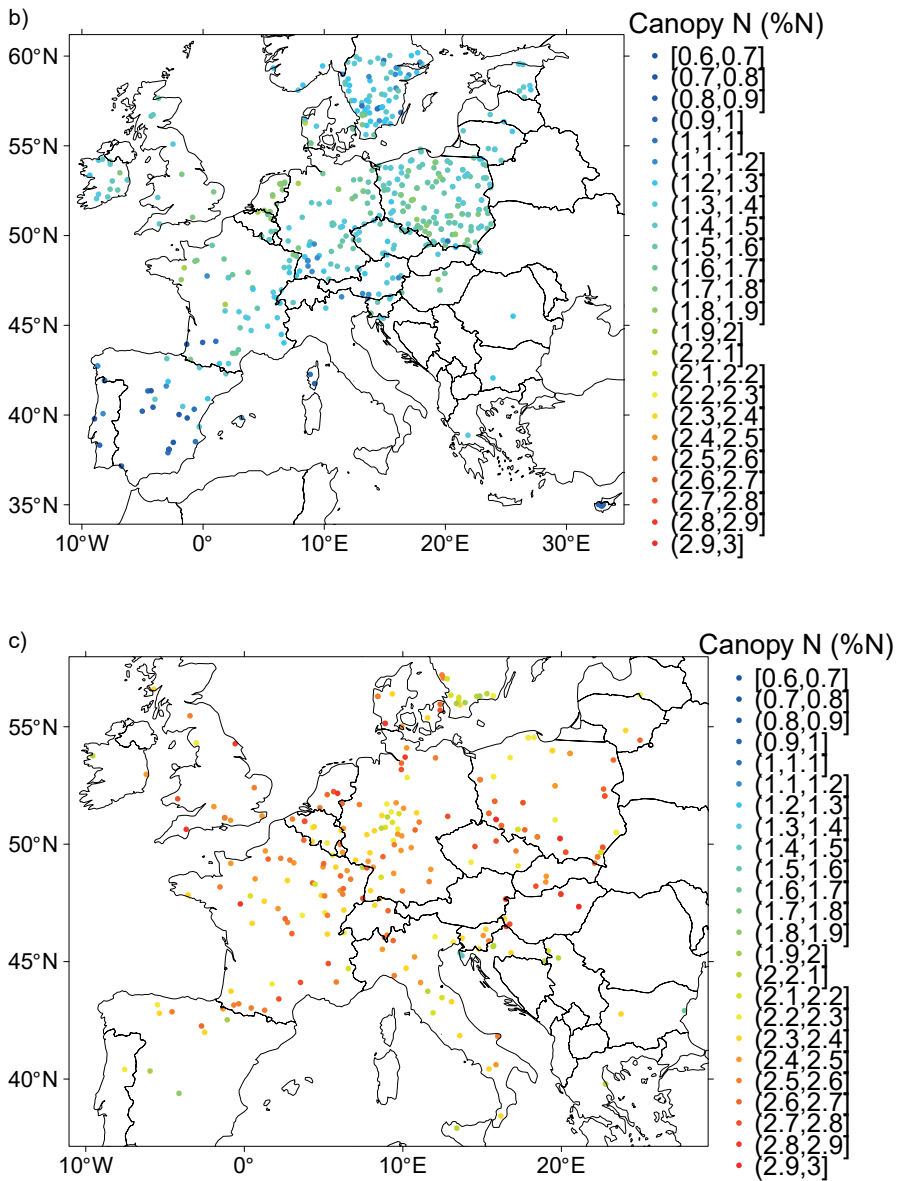


Figure 4.1. Map of the forest plot locations for a) All plots, b) Evergreen Needleleaf Forest (ENF) plots and c) Deciduous Broadleaf Forest (DBF) plots. The color scale gives the long-term canopy nitrogen (%N).

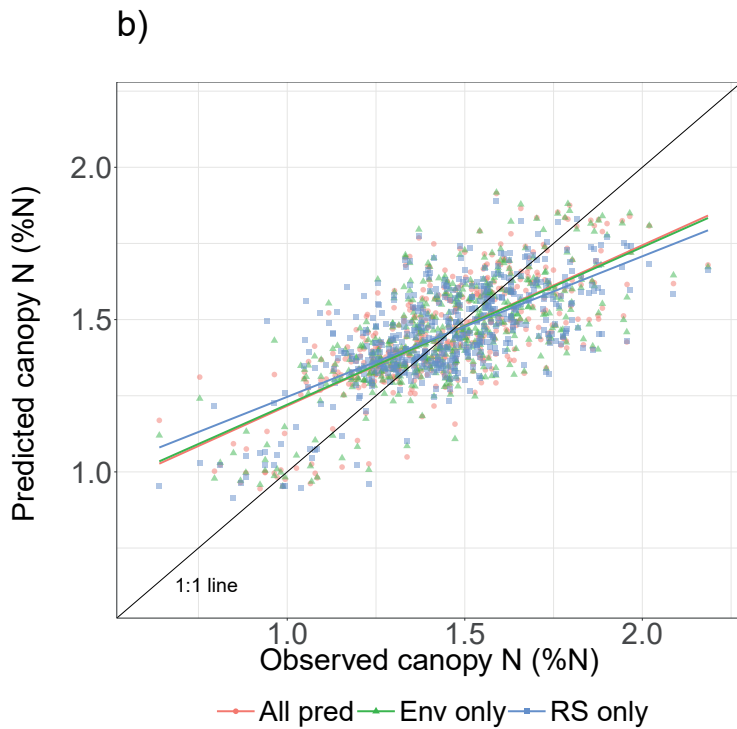
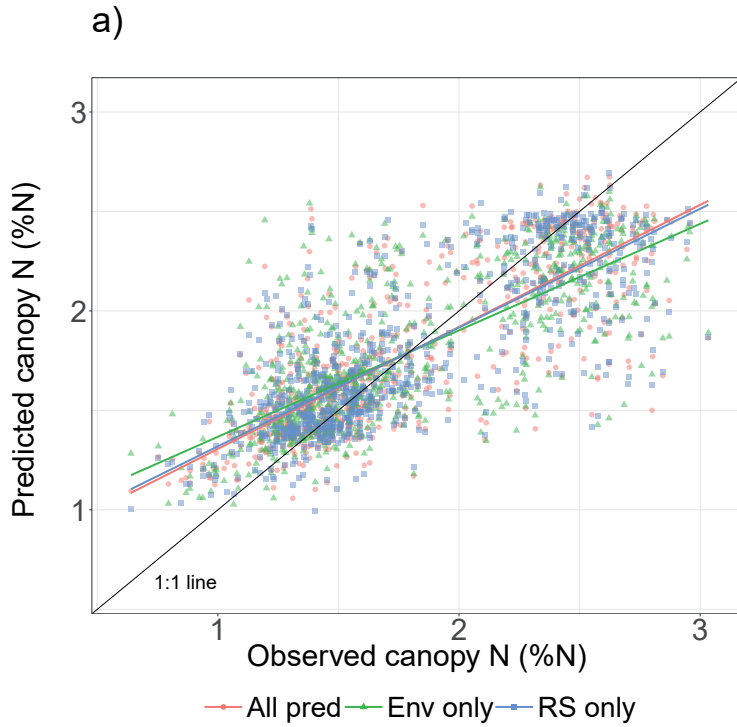
4.3.2 Results of the random forests analysis

Table 4.3 shows results from the random forests analysis. Among the nine models tested, the model including all predictor variables and all plots performed best, with a r^2 of 0.62 for the validation. The models including all predictors showed higher r^2 and lower RRMSE, both for OOB and validation, compared to models including either remote sensing or environmental variables. This was the case for all groups considered (All plots, ENF and DBF). When considering models for either all plots or ENF plots, the type of model (*All pred*, *RS only* or *Env only*) did not have a strong effect on the models fit. For the DBF subgroup, on the contrary, the *RS only* type of model showed lower r^2 compared to the other two ($r^2 = 0.09$ and $r^2 = 0.39$, for *RS only* and both *All pred and Env only*, respectively). The *RS only* model was thus not able to predict canopy N for the DBF subgroup. If we compare the r^2 for different subgroups included in the analysis, the models with all plots performed always better (r^2_{C-Val} 0.54-0.62) always performed better than those of the ENF (r^2_{C-Val} 0.47-0.49) or the DBF (r^2_{C-Val} 0.09-0.39) subgroups. However, the opposite is observed when comparing the RRMSE for three subgroups. The RRMSE for the DBF subgroup (RRMSE 0.08-0.09) is lower than the RRMSE of the ENF (RRMSE 0.11-0.12) or all plots (RRMSE 0.18-0.20) subgroups.

Scatterplots of predicted vs observed values for canopy N for the All plots, ENF plots and DBF plots are presented in Figure 4.2. Regression lines were fitted between the predicted vs observed values for each group of predictor variables studied. Chow tests (Chow 1960) were done to assess whether the sets of coefficients between different linear regressions were equal. The tests showed that the differences between each group-wise pairs of regression lines were not significant.

Table 4.3. Results of the random forests analysis for the models including all predictor variables (*All pred*), only remote sensing variables (*RS only*) and only environmental variables (*Env only*), and for each plant functional type (PFT) group: all plots (*All plots*), Evergreen Needleleaf Forest (*ENF*) and Deciduous Broadleaf Forest (*DBF*). The initial number of predictors represent the number of predictor variables available to build the model before variables selection, the selected number of predictors represents the number of variables selected to build the model. r^2 , Relative Root Mean Squared error (RRMSE) are presented for both out-of-bag data (OOB) and cross-validation (C-Val).

PFT	Number of plots	Model	Initial number of predictors	Selected number of predictors	Calibration (OOB)		Validation (C-Val)	
					r^2	RRMSE	r^2	RRMSE
All plots	818	All pred	86	17	0.63	0.18	0.62	0.18
		RS only	37	16	0.61	0.18	0.60	0.18
		Env only	49	15	0.55	0.19	0.54	0.20
ENF	513	All pred	86	22	0.50	0.11	0.49	0.11
		RS only	37	9	0.45	0.12	0.44	0.12
		Env only	49	8	0.48	0.11	0.47	0.11
DBF	265	All pred	86	17	0.40	0.08	0.39	0.08
		RS only	37	17	0.12	0.09	0.09	0.09
		Env only	49	10	0.39	0.08	0.39	0.08



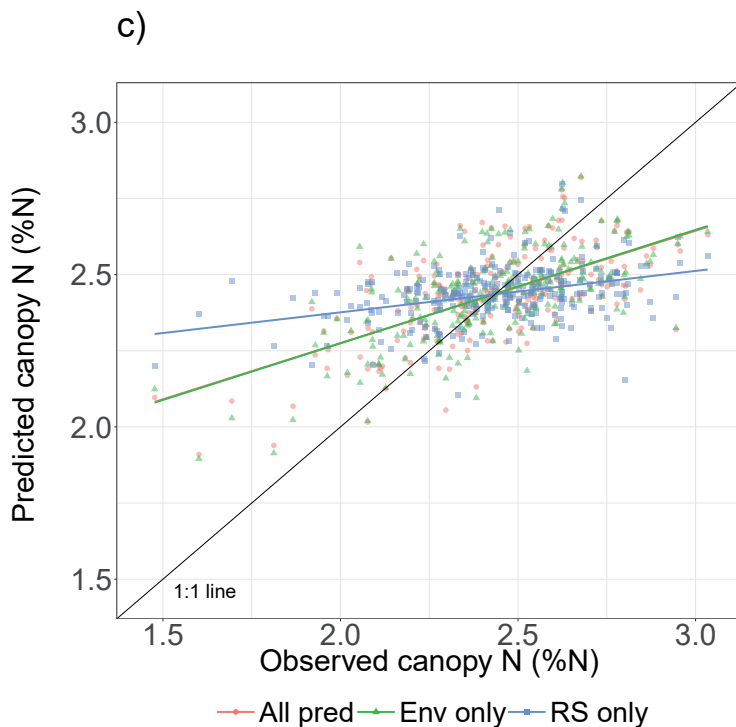


Figure 4.2. Scatterplots between observed and predicted canopy N values (%N) based on OOB prediction for a) All plots, b) ENF plots and c) DBF plots. The color scale refers to the predictor variables tested: all predictor variables (All pred), only remote sensing variables (RS only) and only environmental variables (Env only).

4.3.3 Variable importance

The 10 most important variables for predicting canopy N are presented in Table 4.4 for all models considered. For the *All pred* model, when all plots were included, nine out of the 10 most important variables for predicting canopy N were remote sensing variables, with the two most important ones being EVI long-term average in May and June. These two variables are also the most important for canopy N prediction in the *All plots RS only* model. More generally, the important variables for the *All pred* model showed large agreement with the *RS only* model. In the *Env only* model, the two most important variables were the binary variables for presence or absence of broadleaf deciduous forest and needleleaf evergreen forest. These variables were obtained by modifying the CCI land cover map.

For the PFTs specific models, both *ENF and DBF All pred* models showed large similarities with their *Env only* model counterpart. For the *ENF All pred* model no RS variable was among the 10 most important predictor variables. When comparing *ENF and DBF All pred* models, soil properties and climate variables are important predictors of canopy N. For ENF plots, the soil sand content and the annual mean temperature were more important while for DBF plots, the oxidized nitrogen deposition as well as the mean temperature of the warmest quarter were

Table 4.4. List of the 10 most important variables for canopy nitrogen prediction and importance values for all models tested: all predictor variables (All pred), only remote sensing variables (RS only) and only environmental variables (Env only), and for each plant functional type (PFT) group: All plots, Evergreen Needleleaf Forest (ENF) and Deciduous Broadleaf Forest (DBF).

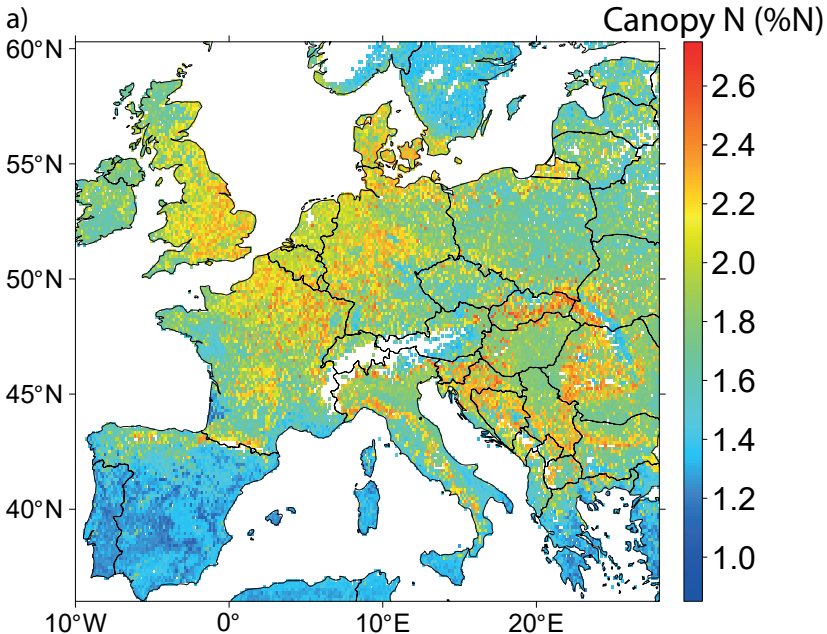
All pred – All plots		Importance	Env only – All plots	Importance	RS only – All plots	Importance
MODIS EVI long-term monthly average May		6.9	CCI land cover broadleaf deciduous	7.9	MODIS EVI long-term monthly average May	6.7
MODIS EVI long-term monthly average June		5.7	CCI land cover needleleaf evergreen	7.4	MODIS EVI long-term monthly average June	5.8
MODIS EVI long-term monthly average July		4.3	Annual Mean Temperature	4.7	MTCI long-term monthly average February	5.7
MTCI long-term monthly average February		4.2	Oxidized nitrogen deposition	4.0	MODIS EVI long-term monthly average July	4.4
MODIS NDVI long-term monthly average June		3.3	Sand content mass fraction	3.7	MODIS EVI long-term monthly average August	3.6
MODIS NDVI long-term monthly average August		2.2	Mean temperature coldest quarter	3.7	MODIS NDVI long-term monthly average June	3.2
MODIS NDVI long-term monthly average November		2.2	Silt content mass fraction	3.2	MODIS NDVI long-term monthly average November	2.5
MTCI long-term monthly average June		2.1	Minimum temperature coldest month	3.0	MTCI long-term monthly average March	2.4
Oxidized nitrogen deposition		1.9	Mean temperature warmest quarter	2.7	MODIS NDVI long-term monthly average August	2.2
MODIS NDVI long-term monthly average March		1.9	Clay content mass fraction	2.6	MTCI long-term monthly average June	2.2
All pred – ENF plots		Importance	Env only – ENF plots	Importance	RS only – ENF plots	Importance
Sand content mass fraction		0.8	Annual mean temperature	1.4	MTCI long-term monthly average February	1.7
Annual mean temperature		0.7	Mean temperature warmest quarter	1.4	MTCI long-term monthly average September	1.2
Coarse fragments volumetric		0.7	Coarse fragments volumetric	1.2	MTCI long-term monthly average March	1.1
Mean temperature warmest quarter		0.6	Precipitation warmest quarter	1.0	MTCI long-term monthly average August	1.1
Precipitation warmest quarter		0.5	Mean temperature coldest quarter	1.0	MTCI long-term monthly average October	0.8
Mean temperature driest quarter		0.5	Mean temperature coldest quarter	0.9	MODIS NDVI long-term monthly average June	0.7
Silt content mass fraction		0.5	Silt content mass fraction	0.7	MTCI long-term monthly average May	0.6
Maximum temperature warmest month		0.3	Temperature seasonality	0.6	MTCI long-term monthly average April	0.5
Oxidized nitrogen deposition		0.3	Soil bulk density	/	MODIS NIR long-term monthly average November	0.4
Mean temperature coldest quarter		0.3	/	/	/	/
All pred – DBF plots		Importance	Env only – DBF plots	Importance	RS only – DBF plots	Importance
Oxidized nitrogen deposition		1.1	Oxidized nitrogen deposition	1.6	MODIS NIR long-term monthly average September	0.8
Mean temperature warmest quarter		0.6	Mean temperature warmest quarter	0.8	MODIS EVI long-term monthly average October	0.8
Sand content mass fraction		0.5	Sand content mass fraction	0.7	MODIS EVI long-term monthly average June	0.7
Silt content mass fraction		0.5	Silt content mass fraction	0.7	MODIS NIR long-term monthly average July	0.7
Maximum temperature warmest month		0.5	Total nitrogen deposition	0.7	MODIS EVI long-term monthly average July	0.6
Annual mean temperature		0.5	Mean temperature driest quarter	0.6	MODIS NIR long-term monthly average October	0.6
Mean temperature driest quarter		0.4	Clay content mass fraction	0.6	MODIS EVI long-term monthly average August	0.5
Total nitrogen deposition		0.4	Soil bulk density	0.5	MTCI long-term monthly average June	0.5
Clay content mass fraction		0.4	Temperature seasonality	0.5	MODIS NDVI long-term monthly average August	0.5
MTCI long-term monthly average May		0.4	Altitude	0.5	MODIS NDVI long-term monthly average October	0.4

important predictors of canopy N. For both DBF and ENF plots, the granulometry of the soil influenced the prediction of canopy N.

4.3.4 Canopy N map for European forests

The best predicted canopy N maps for each group considered, i.e. All plots, ENF plots and DBF plots, are presented in Figure 4.3. The range of canopy N values of the predicted map corresponded to the range observed from forest samplings for each subgroup. The broad-scale spatial patterns show similarities between the three maps considered: in the southern and northern regions of Europe, i.e. the Mediterranean region and the south of Sweden, the predicted canopy N was lower than in the midwestern region of Europe. Local-scale patterns are also present. For example, in Netherlands and in the northwest of Germany, the predicted canopy N is relatively higher than average for both the ENF and DBF maps. However, this is not observed when all plots are considered. The All plots model also predicts relatively higher values of canopy N, approximately between 2.4 and 2.6 %N, in Eastern Europe and Slovakia, in particular. The observed fine-scale pattern corresponds to the location of the Carpathian Mountains. This pattern is also present, although less clearly, in the ENF map, but not in the DBF map.

Figure 4.3



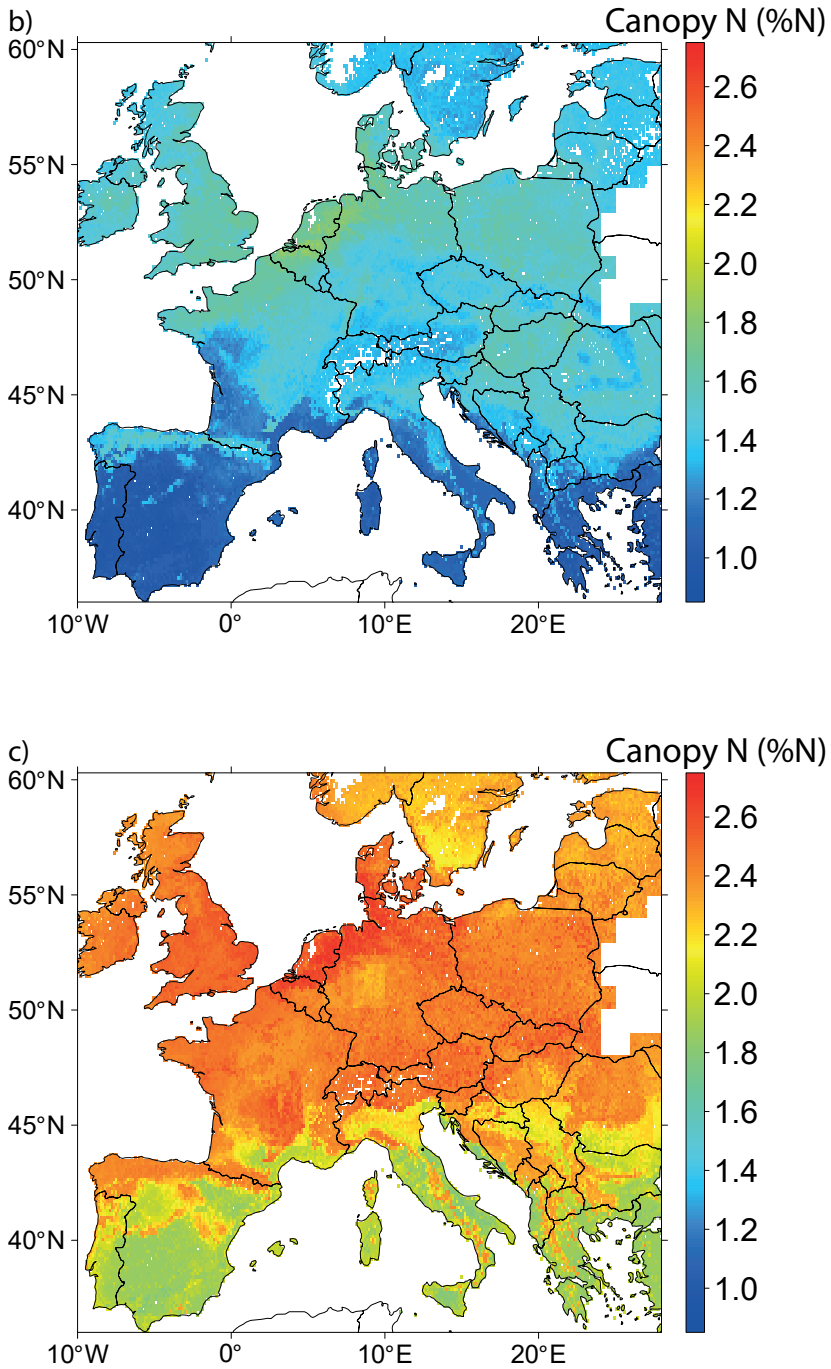


Figure 4.3. Predicted canopy nitrogen maps (%N) calibrated using a) All plots, b) Evergreen Needleleaf Forests (ENF) and c) Deciduous Broadleaf Forest (DBF).

4.4 Discussion

4.4.1 Canopy N spatial pattern

The aim of this study was to estimate spatial patterns in canopy N over European forests, which we were able to do with an r^2 for the validation of 0.62 for the All plots subgroup. The broad-scale spatial pattern (Figure 4.3) of the predicted canopy N maps showed similarities between the three subgroups considered: lower canopy N concentration in the south and in the north of Europe, higher values in mid-western Europe. This pattern was also similar to that observed in the forest plot data (Figure 4.1a), which is expected as this data was used to train the model. This indicates that the developed model was able to represent the broad-scale canopy N pattern present in the data. More specifically, however, the three predicted maps show finer scale variations. Both the DBF and the ENF predicted maps show relatively high values in the Netherlands and in the northwestern part of Germany. The forest plot data (Figure 4.1a) also included several forest plots with relatively high long-term canopy N in this region. This trend, however, is not present in the predicted canopy N map for all plots. Another dissimilarity was noticeable in mid-southern part of France with higher canopy N values in both the All Plots and DBF predicted maps. This region corresponds to the location of the Massif Central mountain area and the canopy N values of forest plots located there were not higher than average.

4.4.2 Comparison with published studies

As mapping canopy N in forests has seldom been done at the European scale, there is not a large body of literature to compare our results with. A recent study published a forest leaf N map at global scale (Moreno-Martínez et al. 2018). When visually comparing this map with the results of the present study for the All plots model, the global scale map also presents relatively lower concentration in the south of Sweden, while showing relatively higher values in the center of Europe. However, the range of values is different between the two maps: while most values are between 1.4 and 2 %N in the published study, in our analysis the range of values is larger, between 1.2 and 2.4 %N. The accuracy for predicting canopy N ($r^2 = 0.62$ for the best model) was similar for both studies. The published study used the published data from the TRY database (Kattge et al. 2011) for calibration. Although the data in the TRY database are numerous, the data were sampled for various research purposes and the sampling methods do not follow a standard guideline. The ICP Forests data, on the contrary, were sampled following a consistent process, which ensures of the good quality of the dataset.

4.4.3 The role of environmental variables

The second aim of this study was to test whether including environmental variables as predictors improves canopy N predictions compared to approaches that rely on remotely sensed data alone. The results showed that including environmental variables as predictors improved the explanatory power of the models for all groups considered as the models for *All pred* always show higher r^2 and lower RRMSE compared to the *Env only* and *RS only* models (Table 4.3). In a previous study in savannah grass, including environmental variables in a stepwise multilinear regression also improved canopy N prediction compared to using VIs only (Ramoelo et al. 2012). It is interesting to note that the *RS only* model performed better

than the *Env only* model for the all plots subgroup, while it was the opposite for both the ENF and DBF subgroups. This shows that, in our study, RS variables were useful to distinguish between different PFTS in the *All pred All plots* model. Moreover, the influence of including environmental variables can also be seen in the most important variables selected for the *All pred* models. For both the *ENF and DBF All pred* models, the most important variables are environmental variables, RS variables being among the least important predictors (Table 4.4). For the DBF subgroup, the *RS only* model showed the lowest observed accuracy of all models tested ($r^2 = 0.09$). The remote sensing variables were not able to predict canopy N and including environmental variables was essential to predict canopy N for DBF plots, in our dataset. Including environmental variables to predict canopy N was thus more beneficial for separate PFTs, and even more so for the DBF subgroup.

4.4.4 Variables importance

Regarding the difference in RS products selected, it is interesting to note that, although the NIR spectral region was shown to be important for canopy N prediction in previous studies (Moreno-Martínez et al. 2018; Mutowo et al. 2018; Ollinger et al. 2008), it was seldom selected as predictor variable in the models tested. It was among the most important variables for the *DBF Env only* model, which showed the lowest r^2 of all the models tested. On the contrary, although being tested in few studies for canopy N prediction (Lepine et al. 2016; Ramoelo et al. 2012; Wang et al. 2012), EVI was the most important variable for the *All pred and all plots* models. A remote sensing product derived from EVI, the maximum EVI, was also found to be the most important variable in a study mapping leaf N at global scale (Moreno-Martínez et al. 2018). In a recent study in the Miombo woodlands, the results showed that NIR VIs, among which NDVI and EVI, are complementary to the NIR spectral region for canopy N mapping (Mutowo et al. 2018). However, in our study, NDVI was not selected as most important variable for any of the developed models. The MTCI index, based on the relationship between canopy N and chlorophyll (Dash and Curran 2004) was only selected as the most important variable for the *ENF RS only* model. The stronger relationship between canopy N and the NIR region compared to the MTCI, a red-edge based index, was previously observed as well. In a mixed temperate forest, Wang et al. (2016) showed that the relationship between canopy N and MTCI was weak. What is also surprising is that the RS products selected as important variables are from different months than the forest plots sampling months. For DBF plots, the two most important RS variables for the *DBF RS only* model were from the months September and October while the DBF plots were mainly sampled in July and August. For the ENF plots, it is not so clear as the RS products from winter months, during which the ENF plots are sampled, were excluded from the analysis.

Among the environmental predictors tested, oxidized N deposition was the most important variable for the DBF plots for both the *All pred and the Env only* model (Table 4.4). For the ENF subgroup, although previous studies showed that canopy N was correlated to N deposition in needleleaf forests (Fleischer et al. 2013; Sardans et al. 2016b), N deposition was only selected among the least important predictors for the *All pred* model and it was not selected at all for the *Env only* model. This might be related to an observed stronger response of deciduous species to N deposition compared to coniferous species (Crowley et al. 2012). The annual mean temperature was among the most important variables for predicting canopy N for the ENF

subgroup for both the *All pred and the Env only* models. Mean annual temperature was also among the selected predictor variables for the *All pred and all plots* model. This is consistent with previous findings showing an influence of mean annual temperature on leaf N (Reich and Oleksyn 2004; Sardans et al. 2015).

For the *Env only* models including all plots, the land cover binary variables indicating the presence of broadleaf deciduous forest and the presence of needleleaf evergreen forest were the most important variables. This is not surprising, as forest type is well known to influence foliage N (Sardans et al. 2016a; Sardans et al. 2015; Sardans et al. 2011). In our study too, the DBF and ENF plots also show different mean canopy N concentration, with 1,4 %N and 2,4 %N for ENF and BDF plots, respectively (Table 4.2).

4.4.5 Source of errors

When we look critically at the results obtained, it is interesting to note that the r^2 values from the 10 fold cross-validation (r^2_{CV}) are close to the r^2_{OOB} values. This shows that the developed models were robust for validation. The r^2 obtained for the ENF and DBF models are lower than the value obtained for the *All pred* models. This could be related to the smaller number of plots used to calibrate the ENF and DBF models, $n = 513$ and $n = 265$, respectively, compared to the *All pred* model, $n = 818$. The limited range of canopy N values for the ENF and DBF models, might also explain the observed decrease in model fit.

4.4.6 Future perspectives

In this study, we showed that combining vegetation indices with environmental variables can contribute to canopy N mapping over large spatial extents. Although, as showed by the comparison and differences with a recent published study (section 4.2), this still needs to be further developed, this study contributes to the discussion about the feasibility of canopy N mapping over large spatial extents. The resulting canopy N map could provide spatial indicators of canopy N in European forests.

In this analysis, we worked with MODIS and MERIS remote sensing data to ensure that the period during which the remote sensing data were acquired was overlapping with the sampling period of the forest plots. Although more recent satellite sensors such as Sentinel 2, Sentinel 3 or RapidEye have either higher spatial or spectral resolutions that would probably improve the accuracy of the obtained canopy N map, the time series were too short to allow for a long enough overlapping period with canopy N sampling data. As the ICP Forest monitoring network is an ongoing project, a future perspective of this study would be to compare the results with a similar analysis, including more recent satellite sensors once the satellites' data time series are long enough.

We focused on canopy N in European forests. While this is a common land use type across Europe (42 % of total land area), it would be valuable to further develop this analysis by including other natural PFTs (like grasses) and non-natural land use types like agricultural land. However, the ICP Forests database we used for this analysis was very valuable, and these high-quality long-term data are not yet available for all land use types. The number of sites, but also the consistency in the way the forest plots are sampled and the %N are measured in ICP Forests is unique. We would like to emphasize that this would be an important necessary step to extrapolate to other land use types.

Finally, an envisioned result of this project is to improve GVMs by providing large-scale information about canopy N and its spatial pattern. In the future, we therefore foresee to compare our results with canopy N modelled with GVMs. However, canopy N values are not static over time, and in our study we averaged the canopy plot data as well as the RS data over a long time period. This gave us more data to work with, as for each year much less data was available. However, if we want to optimize GVMs using the predicted maps, including temporal variations, e.g. yearly or bi-yearly data, would make the predicted maps more compatible with GVMs output. Another future development of the present study is thus to include temporal changes in canopy N values.

4.5 Conclusion

In this study, our objective was to characterize spatial patterns of canopy N in forests across Europe. We showed that we could map canopy N using the random forests technique and calibration data from ICP Forests with good accuracy ($r^2 = 0.62$, RRMSE = 0.18, for validation). Among the RS products tested (EVI, MTCI, NIR and NDVI), EVI was the most important predictor for canopy N prediction when all plots were included, while MTCI the most important predictor for the *ENF RS only* model. We also investigated whether including environmental variables as predictors would improve the prediction models. For all subgroups tested (All plots, ENF plots and DBF plots), including environmental variables improved the predictions. Moreover, in our dataset, including environmental variables was especially essential for the DBF plots, as the prediction model based on remotely sensed data products only was not able to predict canopy N with sufficient accuracy. Finally, including environmental variables together with RS products to predict canopy N showed promising results and could be tested in other regions and with different land use types. A future outcome of this analysis is to compare the predicted canopy N map to GVMs outputs.

4.6 Acknowledgements

This work was supported by the Netherlands Organization for Scientific Research (NWO) [NWO ALW-GO-AO/14-12]. We would like to acknowledge Álvaro Moreno-Martínez for sharing his high-resolution leaf nitrogen map with us as well as Gerard Heuvelink for his help with the random forests model.

4.7 References

- Adjorlolo, C., Mutanga, O., & Cho, M.A. (2014). Estimation of canopy nitrogen concentration across c3 and c4 grasslands using worldview-2 multispectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7, 4385-4392. 10.1109/JSTARS.2014.2320601
- AppEEARS Team (2019). *Application for Extracting and Exploring Analysis Ready Samples (AppEEARS)*, Version 2.21. USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, USA: NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC). Accessed 30 May 2019. <https://lpdaacsvc.cr.usgs.gov/appeears>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32. 10.1023/a:1010933404324
- Brungard, C.W., Boettinger, J.L., Duniway, M.C., Wills, S.A., & Edwards, T.C. (2015). Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*, 239-240, 68-83. <https://doi.org/10.1016/j.geoderma.2014.09.019>
- Chapin, F.S.I. (1987). Adaptations and physiological responses of wild plants to nutrient stress. In W.H. Gabelman, & B.C. Loughman (Eds.), *Genetic Aspects of Plant Mineral Nutrition. Developments in Plant and Soil Sciences* Dordrecht: Springer. https://doi.org/10.1007/978-94-009-3581-5_2.
- Chemura, A., Mutanga, O., Odindi, J., & Kutuywayo, D. (2018). Mapping spatial variability of foliar nitrogen in coffee (*Coffea arabica* L.) plantations with multispectral Sentinel-2 MSI data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 138, 1-11. <https://doi.org/10.1016/j.isprsjprs.2018.02.004>
- Chen, P., Haboudane, D., Tremblay, N., Wang, J., Vigneault, P., & Li, B. (2010). New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sensing of Environment*, 114, 1987-1997. <https://doi.org/10.1016/j.rse.2010.04.006>
- Cho, M.A., Ramoelo, A., Debba, P., Mutanga, O., Mathieu, R., van Deventer, H., & Ndlovu, N. (2013). Assessing the effects of subtropical forest fragmentation on leaf nitrogen distribution using remote sensing data. *Landscape Ecology*, 28, 1479-1491. doi:10.1007/s10980-013-9908-7
- Chow, G.C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, 28, 591-605. 10.2307/1910133
- Clevers, J.G.P.W., & Gitelson, A.A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and -3. *International Journal of Applied Earth Observation and Geoinformation*, 23, 344-351. doi:10.1016/j.jag.2012.10.008
- Crowley, K.F., McNeil, B.E., Lovett, G.M., Canham, C.D., Driscoll, C.T., Rustad, L.E., Denny, E., Hallett, R.A., Arthur, M.A., Boggs, J.L., Goodale, C.L., Kahl, J.S., McNulty, S.G., Ollinger, S.V., Pardo, L.H., Schaberg, P.G., Stoddard, J.L., Weand, M.P., & Weathers, K.C. (2012). Do Nutrient Limitation Patterns Shift from Nitrogen Toward Phosphorus with Increasing Nitrogen Deposition Across the Northeastern United States? *Ecosystems*, 15, 940-957. 10.1007/s10021-012-9550-2
- Dash, J., & Curran, P.J. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25, 5403-5413. doi:10.1080/0143116042000274015
- Defourny, P., Kirches, G., Brockmann, C., Boettcher, M., Peters, M., Bontemps, S., Lamarche, C., Schlerf, M., & Santoro, M. (2016). *Land Cover CCI Product User Guide Version 2.5* (p. 91). Louvain-La-Neuve, Belgium: UCL-Geomatics. Accessed 30 May 2019. <http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-PUG-v2.5.pdf>
- Didan, K. (2015). *MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006* [Dataset]. USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dako-

- ta: NASA EOSDIS Land Processes DAAC. Accessed 28 January 2019. <https://lpdaac.usgs.gov>. <https://doi.org/10.5067/MODIS/MOD13Q1.006>
- European Environment Agency (2013). *Digital Elevation Model over Europe (EU-DEM)* [Dataset]. European Environment Agency. Accessed 20 April 2015. <https://www.eea.europa.eu/data-and-maps/data/eu-dem>
- Evans, J.R. (1989). Photosynthesis and nitrogen relationships in leaves of C3 plants. *Oecologia*, 78, 9-19. doi:10.1007/BF00377192
- Ferreti, M., Fischer, R., Mues, V., Granke, O., Lorenz, M., & Seidling, W. (2017). Part II: Basic design principles for the ICP Forests Monitoring Networks. In UNECE ICP Forests Programme Co-ordinating Centre (Ed.), *Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests* (p. 21 + Annex). Thünen Institute of Forest Ecosystems, Eberswalde, Germany. <http://www.icp-forests.net/page/icp-forests-manual>. ISBN: 978-3-86576-162-0
- Fick, S.E., & Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302-4315. doi:10.1002/joc.5086
- Fleischer, K., Rebel, K.T., Van Der Molen, M.K., Erismann, J.W., Wassen, M.J., Van Loon, E.E., Montagnani, L., Gough, C.M., Herbst, M., Janssens, I.A., Gianelle, D., & Dolman, A.J. (2013). The contribution of nitrogen deposition to the photosynthetic capacity of forests. *Global Biogeochemical Cycles*, 27, 187-199. doi:10.1002/gbc.20026
- Han, W.X., Fang, J.Y., Reich, P.B., Ian Woodward, F., & Wang, Z.H. (2011). Biogeography and variability of eleven mineral elements in plant leaves across gradients of climate, soil and plant functional type in China. *Ecology Letters*, 14, 788-796. doi:10.1111/j.1461-0248.2011.01641.x
- Hansen, P.M., & Schjoerring, J.K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86, 542-553. doi:10.1016/S0034-4257(03)00131-7
- Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangquan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, 12, e0169748. doi:10.1371/journal.pone.0169748
- Hijmans, R.J. (2018). *raster: Geographic Data Analysis and Modeling*, R package version 2.8-4. Accessed 26 April 2019. <https://CRAN.R-project.org/package=raster>
- Hikosaka, K. (2004). Interspecific difference in the photosynthesis-nitrogen relationship: Patterns, physiological causes, and ecological importance. *Journal of Plant Research*, 117, 481-494. doi:10.1007/s10265-004-0174-2
- Homolová, L., Malenovský, Z., Clevers, J.G.P.W., García-Santos, G., & Schaepman, M.E. (2013). Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity*, 15, 1-16. doi:10.1016/j.ecocom.2013.06.003
- Horler, D.N.H., Dockray, M., & Barber, J. (1983). The red edge of plant leaf reflectance. *International Journal of Remote Sensing*, 4, 273-288. doi:10.1080/01431168308948546
- ISIMIP (2019). *Input Data Set: Nitrogen Deposition* [Dataset]. Accessed 27 May 2019. <https://www.isimip.org/gettingstarted/details/24/>
- Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönsch, G., Garnier, E., Westoby, M., Reich, P.B., Wright, I.J., Cornelissen, J.H.C., Violle, C., Harrison, S.P., Van Bodegom, P.M., Reichstein, M., Enquist, B.J., Soudzilovskaia, N.A., Ackerly, D.D., Anand, M., Atkin, O., Bahn, M., Baker, T.R.,

- Baldocchi, D., Bekker, R., Blanco, C.C., Blonder, B., Bond, W.J., Bradstock, R., Bunker, D.E., Casanoves, F., Cavender-bares, J., Chambers, J.Q., Chapin iii, F.S., Chave, J., Coomes, D., Cornwell, W.K., Craine, J.M., Dobrin, B.H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W.F., Fang, J., Fernández-méndez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G.T., Fyllas, N.M., Gallagher, R.V., Green, W.A., Gutierrez, A.G., Hickler, T., Higgins, S.I., Hodgson, J.G., Jalili, A., Jansen, S., Joly, C.A., Kerkhoff, A.J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J.M.H., Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T.D., Leishman, M., Lens, F., Lenz, T., Lewis, S.L., Lloyd, J., Llusà, J., Louault, F., Ma, S., Mahecha, M.D., Manning, P., Massad, T., Medlyn, B.E., Messier, J., Moles, A.T., Müller, S.C., Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V.G., Onoda, Y., Ordoñez, J., Overbeck, G., Ozinga, W.A., Patiño, S., Paula, S., Pausas, J.G., Peñuelas, J., Phillips, O.L., Pillar, V., Poorter, H., Poorter, L., Poschold, P., Prinzing, A., Proulx, R., Rammig, A., Reinsch, S., Reu, B., Sack, L., Salgado-negret, B., Sardans, J., Shiodera, S., Shipley, B., Siefert, A., Sosinski, E., Soussana, J.-F., Swaine, E., Swenson, N., Thompson, K., Thornton, P., Waldram, M., Weiher, E., White, M., White, S., Wright, S.J., Yguel, B., Zaehle, S., Zanne, A.E., & Wirth, C. (2011). TRY – a global database of plant traits. *Global Change Biology*, *17*, 2905-2935. 10.1111/j.1365-2486.2011.02451.x
- Kergoat, L., Lafont, S., Arneth, A., Le Dantec, V., & Saugier, B. (2008). Nitrogen controls plant canopy light-use efficiency in temperate and boreal ecosystems. *Journal of Geophysical Research: Biogeosciences*, *113*. doi:10.1029/2007JG000676
- Kokaly, R.F., Asner, G.P., Ollinger, S.V., Martin, M.E., & Wessman, C.A. (2009). Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment*, *113*, S78-S91. 10.1016/j.rse.2008.10.018
- Kuhn, M. (2018). *caret: Classification and Regression Training*. R package version 6.0-81. Accessed 26 April 2019. <https://CRAN.R-project.org/package=caret>
- Kumar, L., Schmidt, K., Dury, S., & Skidmore, A. (2006). Imaging Spectrometry and Vegetation Science. In F.D.v.d. Meer, & S.M. de Jong (Eds.), *Imaging Spectrometry: Basic Principles and Prospective Applications* (pp. 111-155). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-0-306-47578-8_5. ISBN: 978-0-306-47578-8
- Lamarque, J.F., Dentener, F., McConnell, J., Ro, C.U., Shaw, M., Vet, R., Bergmann, D., Cameron-Smith, P., Dalsoren, S., Doherty, R., Faluvegi, G., Ghan, S.J., Josse, B., Lee, Y.H., MacKenzie, I.A., Plummer, D., Shindell, D.T., Skeie, R.B., Stevenson, D.S., Strode, S., Zeng, G., Curran, M., Dahl-Jensen, D., Das, S., Fritzsche, D., & Nolan, M. (2013a). Multi-model mean nitrogen and sulfur deposition from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): evaluation of historical and projected future changes. *Atmos. Chem. Phys.*, *13*, 7997-8018. 10.5194/acp-13-7997-2013
- Lamarque, J.F., Shindell, D.T., Josse, B., Young, P.J., Cionni, I., Eyring, V., Bergmann, D., Cameron-Smith, P., Collins, W.J., Doherty, R., Dalsoren, S., Faluvegi, G., Folberth, G., Ghan, S.J., Horowitz, L.W., Lee, Y.H., MacKenzie, I.A., Nagashima, T., Naik, V., Plummer, D., Righi, M., Rumbold, S.T., Schulz, M., Skeie, R.B., Stevenson, D.S., Strode, S., Sudo, K., Szopa, S., Voulgarakis, A., & Zeng, G. (2013b). The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics. *Geosci. Model Dev.*, *6*, 179-206. 10.5194/gmd-6-179-2013
- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., & Martin, M.E. (2016). Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sensing of Environment*, *173*, 174-186. doi:10.1016/j.rse.2015.11.028

- Li, F., Miao, Y., Feng, G., Yuan, F., Yue, S., Gao, X., Liu, Y., Liu, B., Ustin, S.L., & Chen, X. (2014). Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Research*, 157, 111-123. doi:10.1016/j.fcr.2013.12.018
- Liaw, A., & Wiener, M. (2002). *R news*, 2, 18-22 <https://CRAN.R-project.org/doc/Rnews/>
- Ling, B., Goodin, D.G., Mohler, R.L., Laws, A.N., & Joern, A. (2014). Estimating canopy nitrogen content in a heterogeneous grassland with varying fire and grazing treatments: Konza Prairie, Kansas, USA. *Remote Sensing*, 6, 4430-4453. 10.3390/rs6054430
- Loozen, Y., Rebel, K.T., Karssenbergh, D., Wassen, M.J., Sardans, J., Peñuelas, J., & de Jong, S.M. (2018). Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index. *Biogeosciences*, 15, 2723-2742. 10.5194/bg-15-2723-2018
- Martin, M.E., Plourde, L.C., Ollinger, S.V., Smith, M.L., & McNeil, B.E. (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment*, 112, 3511-3519. doi:10.1016/j.rse.2008.04.008
- McNeil, B.E., Read, J.M., & Driscoll, C.T. (2007). Foliar nitrogen responses to elevated atmospheric nitrogen deposition in nine temperate forest canopy species. *Environmental Science and Technology*, 41, 5191-5197. doi: 10.1021/es062901z
- McNeil, B.E., Read, J.M., & Driscoll, C.T. (2012). Foliar Nitrogen Responses to the Environmental Gradient Matrix of the Adirondack Park, New York. *Annals of the Association of American Geographers*, 102, 1-16. 10.1080/00045608.2011.595654
- 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. doi:10.2111/04-17.1
- Moreno-Martínez, Á., Camps-Valls, G., Kattge, J., Robinson, N., Reichstein, M., van Bodegom, P., Kramer, K., Cornelissen, J.H.C., Reich, P., Bahn, M., Niinemets, Ü., Peñuelas, J., Craine, J.M., Cerabolini, B.E.L., Minden, V., Laughlin, D.C., Sack, L., Allred, B., Baraloto, C., Byun, C., Soudzilovskaia, N.A., & Running, S.W. (2018). A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sensing of Environment*, 218, 69-88. <https://doi.org/10.1016/j.rse.2018.09.006>
- Mutanga, O., Adam, E., Adjorloloa, C., & Abdel-Rahmanw, E.M. (2015). Evaluating the robustness of models developed from field spectral data in predicting African grass foliar nitrogen concentration using WorldView-2 image as an independent test dataset. *International Journal of Applied Earth Observation and Geoinformation*, 34, 178-187. 10.1016/j.jag.2014.08.008
- 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. 10.1016/j.rse.2003.11.001
- Mutowo, G., Mutanga, O., & Masocha, M. (2018). Evaluating the Applications of the Near-Infrared Region in Mapping Foliar N in the Miombo Woodlands. *Remote Sensing*, 10. 10.3390/rs10040505
- NEODC (2015). *NEODC-NERC Earth Observation Data Centre* [Webpage]. UK: Natural Environment Research Council. Accessed 6 February 2015. <http://neodc.nerc.ac.uk/>
- Ollinger, S.V., Richardson, A.D., Martin, M.E., Hollinger, D.Y., Frolking, S.E., Reich, P.B., Plourde, L.C., Katul, G.G., Munger, J.W., Oren, R., Smith, M.L., Paw U, K.T., Bolsta, P.V., Cook, B.D., Day, M.C., Martin, T.A., Monson, R.K., & 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 Sciences of the United States of America*, 105, 19336-19341. doi:10.1073/pnas.0810021105.

- R Core Team (2019). *R: A language and environment for statistical computing* [Computer program]. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramoelo, A., Cho, M.A., Mathieu, R., Madonsela, S., van de Kerchove, R., Kaszta, Z., & Wolff, E. (2015). Monitoring grass nutrients and biomass as indicators of rangeland quality and quantity using random forest modelling and WorldView-2 data. *International Journal of Applied Earth Observation and Geoinformation*, 43, 43-54. <https://doi.org/10.1016/j.jag.2014.12.010>
- Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., & Heitkönig, I.M.A. (2012). Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor. *International Journal of Applied Earth Observation and Geoinformation*, 19, 151-162. doi:10.1016/j.jag.2012.05.009
- Rautio, P., Fürst, A., Stefan, K., Raitio, H., & Bartels, U. (2016). Sampling and Analysis of Needles and Leaves. In UNECE ICP Forests Programme Co-ordinating Centre (Ed.), *Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests* (p. 19 + Annex). Thünen Institute of Forest Ecosystems, Eberswalde, Germany. <http://www.icp-forests.org/manual.htm>. ISBN: 978-3-86576-162-0
- Reich, P.B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society B: Biological Sciences*, 279, 2128-2134. doi:10.1098/rspb.2011.2270
- Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Volin, J.C., & Bowman, W.D. (1999). Generality of leaf trait relationships: A test across six biomes. *Ecology*, 80, 1955-1969. doi:10.2307/176671
- 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 of the United States of America*, 101, 11001. 10.1073/pnas.0403588101
- Sardans, J., Alonso, R., Carnicer, J., Fernández-Martínez, M., Vivanco, M.G., & Peñuelas, J. (2016a). Factors influencing the foliar elemental composition and stoichiometry in forest trees in Spain. *Perspectives in Plant Ecology, Evolution and Systematics*, 18, 52-69. 10.1016/j.ppees.2016.01.001
- Sardans, J., Alonso, R., Janssens, I.A., Carnicer, J., Veresoglou, S., Rillig, M.C., Fernández-Martínez, M., Sanders, T.G.M., & Peñuelas, J. (2016b). Foliar and soil concentrations and stoichiometry of nitrogen and phosphorous across European *Pinus sylvestris* forests: Relationships with climate, N deposition and tree growth. *Functional Ecology*, 30, 676-689. 10.1111/1365-2435.12541
- Sardans, J., Janssens, I.A., Alonso, R., Veresoglou, S.D., Rillig, M.C., Sanders, T.G.M., Carnicer, J., Filella, I., Farré-Armengol, G., & Peñuelas, J. (2015). Foliar elemental composition of European forest tree species associated with evolutionary traits and present environmental and competitive conditions. *Global Ecology and Biogeography*, 24, 240-255. 10.1111/geb.12253
- Sardans, J., Rivas-Ubach, A., & Peñuelas, J. (2011). Factors affecting nutrient concentration and stoichiometry of forest trees in Catalonia (NE Spain). *Forest Ecology and Management*, 262, 2024-2034. doi:10.1016/j.foreco.2011.08.019
- Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, 25, 47-54. doi:10.1016/j.jag.2013.04.003
- Serrano, L., Peñuelas, J., & Ustin, S.L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, 81, 355-364. doi:10.1016/S0034-4257(02)00011-1

- Tian, Y.C., Yao, X., Yang, J., Cao, W.X., Hannaway, D.B., & Zhu, Y. (2011). Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance. *Field Crops Research*, *120*, 299-310. doi:10.1016/j.fcr.2010.11.002
- Wang, W., Yao, X., Yao, X., Tian, Y., Liu, X., Ni, J., Cao, W., & Zhu, Y. (2012). Estimating leaf nitrogen concentration with three-band vegetation indices in rice and wheat. *Field Crops Research*, *129*, 90-98. <http://dx.doi.org/10.1016/j.fcr.2012.01.014>
- Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A.K., Jones, S., Suarez, L., Woodgate, W., Heiden, U., Heurich, M., & Hearne, J. (2016). Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sensing*, *8*. doi:10.3390/rs8060491
- Wright, I.J., Reich, P.B., Cornelissen, J.H.C., Falster, D.S., Garnier, E., Hikosaka, K., Lamont, B.B., Lee, W., Oleksyn, J., Osada, N., Poorter, H., Villar, R., Warton, D.I., & Westoby, M. (2005). Assessing the generality of global leaf trait relationships. *New Phytologist*, *166*, 485-496. 10.1111/j.1469-8137.2005.01349.x
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., & Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, *428*, 821-827. doi:10.1038/nature02403
- Xu-Ri, & Prentice, I.C. (2008). Terrestrial nitrogen cycle simulation with a dynamic global vegetation model. *Global Change Biology*, *14*, 1745-1764. 10.1111/j.1365-2486.2008.01625.x

Chapter 5

Comparison of simulated foliage
nitrogen by the O-CN and
LPJ-GUESS vegetation models
with a canopy nitrogen map
based on forest sampling



Comparison of simulated foliage nitrogen by the O-CN and LPJ-GUESS vegetation models with a canopy nitrogen map based on forest sampling

Abstract. Terrestrial carbon (C) storage is simulated using global vegetation models (GVMs). Nitrogen (N) availability influences C assimilation by the terrestrial biosphere. To account for this link between the N and C cycles, GVMs include an N cycle module to simulate the effect of N availability on terrestrial C storage. Compared to data on the C cycle that are globally available, data on the N cycle are however lacking at the global scale. In order to calibrate and validate the N cycle modules of GVMs, there is a need to evaluate the models against N observations. Recently, earth observation data combined with ancillary datasets enabled mapping of foliage N over large areas. The availability of this data facilitates the comparison of the spatial pattern of canopy N simulated by the GVMs. In this study, we compared the spatial patterns of canopy N from two GVMs, O-CN and LPJ-GUESS, with a canopy N map, calibrated using forest plot foliar N data in Europe using a random forests model, and with each other. The results showed that the magnitude and spatial patterns of canopy N values predicted by both GVMs was comparable to those from the random forests model. In the northern region of Europe, the three models predicted similar spatial patterns. The canopy N values predicted by the models showed significant linear relationships in northern Europe. In the south of Europe, in the Mediterranean region, however, the vegetation models predicted contrasting spatial patterns with the random forests map and with each other. The results showed that for the northern region of Europe, while the physiological processes are represented differently in the two GVMs, the results obtained regarding canopy N are comparable and the N cycle representation in the models are coherent with observations. No clear reason for the different behavior for the Mediterranean part surfaced so far. Future research is needed to understand the cause of the divergent spatial patterns observed in the southern region of Europe.

5.1 Introduction

The atmospheric CO₂ concentration has risen globally since pre-industrial time due to anthropogenic CO₂ emissions most probably causing climate change (Bindoff et al. 2013; Friedlingstein et al. 2019). During 2009-2018, approximately 29% of the CO₂ emitted annually, from fossil fuel emissions and land use change, was estimated to be stored in terrestrial vegetation, which reduced the increase in CO₂ concentration in the atmosphere (Friedlingstein et al. 2019). Forests and their soils are important carbon (C) sinks, accounting for up to 65 % of terrestrial organic C (Reichstein and Carvalhais 2019). The fate and magnitude of the terrestrial C sink is however uncertain with future global change (Ahlström et al. 2012; Ahlström et al. 2017; Huntzinger et al. 2017; Lovenduski and Bonan 2017; Meyerholt et al. 2016; Schurgers et al. 2018).

Global vegetation models (GVMs) are used to estimate terrestrial C storage, at present times and under future conditions (Friedlingstein et al. 2019). GVMs simulate vegetation growth and mortality as well as the associated C, water and energy fluxes between the soil, the atmosphere and the vegetation. Using GVMs, it is possible to simulate vegetation C uptake under rising global temperature and atmospheric CO₂ concentration (Ahlström et al. 2012; Ahlström et al. 2017). The influence of nutrient availability on the C cycle is also studied (Wieder et al. 2015).

Nitrogen (N) is an essential nutrient for vegetation growth (Chapin 1987). N availability poses a constrain on C sequestration by terrestrial ecosystems (Fernández-Martínez et al. 2014; LeBauer and Treseder 2008; Vitousek and Howarth 1991). To account for this link between the C and N cycle, N cycle modules are included in GVMs (Smith et al. 2014; Zaehle and Dalmonech 2011; Zaehle and Friend 2010; Zaehle et al. 2014). By including the N cycle, GVMs enable to project the combined influence of atmospheric warming, N deposition and increasing atmospheric CO₂ on terrestrial C storage (Wårlind et al. 2014; Zaehle 2013; Zaehle et al. 2010). These drivers have contrasting effects on the predicted C storage. While a higher CO₂ concentration increases terrestrial C assimilation through the CO₂ fertilization effect, including the influence of N availability mostly reduces the fertilization effect previously observed in GVMs that do not incorporate the link between C and N cycles (Huntzinger et al. 2017; Wieder et al. 2015). However, uncertainties remain regarding the magnitude and direction of the influence of the N cycle on the C assimilation under future change. One GVM in particular, LPJ-GUESS, even predicts an increase in C storage by 2100 when taking the N cycle into account (Wårlind et al. 2014).

GVMs with N cycle modules are process-based, i.e. vegetation growth and mortality are simulated based on known physiological processes. Observed physiological relationships between foliar N concentration and photosynthesis (Reich et al. 1999; Wright et al. 2004) are incorporated in the models. Leaf N is an important variable in the models as both gross photosynthesis and respiration are often functions of leaf N (Smith et al. 2014; Zaehle and Friend 2010; Zaehle et al. 2014). Leaf N concentration is important for other biomass compartments as well, as N allocation to these compartments are prescribed relative to leaf N (Smith et al. 2014; Zaehle and Friend 2010). More generally, the vegetation N status also influences the simulated whole plant growth as well as C allocation between leaves and roots (Zaehle et al. 2014). Given that leaf N has a central role in the simulated C assimilation by

vegetation, it is necessary to evaluate if the models accurately predict foliage N. Evaluation of the predicted leaf N has already been done for specific models. For example, the O-CN and LPJ-GUESS models were able to simulate realistic leaf N values at specific sites (Fleischer et al. 2015; Zaehle and Friend 2010). However, although leaf N is an essential variable within these models, no evaluation of the spatial pattern of the predicted foliage N has been done at large scale due to the lack of a suitable large scale and high-quality dataset (Zaehle and Dalmonech 2011).

In this context, data available through earth observation represents an opportunity for the evaluation of GVMs. Remote sensing has already proven useful to study vegetation dynamics and the C cycle (Exbrayat et al. 2019). Remote sensing techniques have also been developed to detect canopy N at various scales (Loozen et al. 2019; Loozen et al. 2020; Loozen et al. 2018). Initially used to sense leaf N at small scale in crops using spectroradiometers (Hansen and Schjoerring 2003; Xue et al. 2004), remote sensing of canopy N has expanded to various vegetation types, among which forests (Huber et al. 2008; Majeke et al. 2008; McNeil et al. 2007; Schlerf et al. 2010). Different analytical techniques applied to spaceborne imagery have been developed to map canopy N at regional scales (Loozen et al. 2018; Lepine et al. 2016; Ramoelo et al. 2012). Recently, the combination of both data from satellite sensors and environmental variables with machine learning techniques has enabled mapping canopy N at large scales (Loozen et al. 2020; Moreno-Martínez et al. 2018).

Here, we evaluate the magnitude and spatial patterns of the long-term average foliage N modelled by two GVMs that include an N cycle module. To do so, we compare foliage N maps predicted by the two GVMs with a canopy N map of European forest created with extensive forest sampling data, remote sensing and environmental variables (Loozen et al. 2020). The vegetation models we compare are the land surface model O-CN (Zaehle and Friend 2010) and the dynamic vegetation model LPJ-GUESS (Smith et al. 2014). We chose to include these vegetation models because they are widely-used and differ in the C and N cycling process representation. The canopy N map is based on a random forests model, which included both remote sensing and environmental variables as predictors. The random forests model was based on canopy N data from sampled forest plots in Europe. Our objective is to evaluate the two GVMs and the random forests model with regards to a) the magnitude of the differences in the modelled canopy N and b) the differences in the spatial patterns of modelled canopy N.

5.2 Material and methods

5.2.1 Global vegetation models

5.2.1.1 O-CN

O-CN is a land surface model based on the ORCHIDEE model (Zaehle and Friend 2010). It includes energy, water and carbon fluxes as well as an N cycle module. The base unit in the O-CN model is an average representative individual of a plant functional type (PFT) on a grid cell. The average individual is represented by different biomass compartments between which fluxes of C, N, water and energy occur. The C and N fluxes between the different compartments are calculated on a half-hourly timestep while vegetation growth and population dynamics are simulated daily. The vertical distribution of leaf N within the canopy

is represented by dividing the canopy in distinct layers for which the photosynthetic processes are calculated separately. The photosynthesis processes, i.e. the maximum carboxylation and electron transport, are proportional to the fraction of leaf N associated with the photosynthetic apparatus in each layer (Zaehle and Friend 2010). Autotrophic respiration is dependent on tissue N concentration and temperature. Plant growth is a balance of the labile C pool, gross photosynthesis and autotrophic respiration and is limited by the availability of N in the labile N pool. Plant N uptake is proportional to the plant N status, fine root mass, the availability of soil mineral N and soil temperature. Further explanation of the O-CN model can be found in Zaehle and Friend (2010) and references therein. The O-CN model was run for the period 1980-2017. The simulation for the O-CN model included historical changes in climate, atmospheric CO₂ concentration, land-use and N deposition. The climate forcing included temperature, precipitation and incoming radiation from the merged monthly CRU (Harris et al. 2014) and 6-hourly JRA-55 dataset (Kobayashi et al. 2015), updated to 2017 following the methodology in (Viovy 2018). The model was also forced with global atmospheric CO₂ concentration (Dlugokencky and Tans 2018) and historical N deposition (Hegglin et al. 2016). Historical land-use changes were obtained from an update of the harmonized land-use dataset (Hurtt et al. 2011). The spatial resolution of the O-CN map was 1 degree.

5.2.1.2 LPJ-GUESS

LPJ-GUESS, which stands for Lund-Potsdam-Jena General Ecosystem Simulator, is a process-based dynamic vegetation model, which simulates competition for resources between different vegetation patches on a grid cell (Smith et al. 2014). The patches are characterized by different succession stages. Physiological processes, including photosynthesis, respiration, stomatal conductance and phenology, are calculated daily while C allocation to each plant biomass compartment (leaves, fine roots and sapwood) from newly increased net primary production (NPP) is simulated yearly. Plants are assumed to take up N from the soil mineral N pool, which is supplied by N atmospheric deposition, biological N fixation and N mineralization. Plants gross photosynthesis depends on leaf N and leaf N demand is thus calculated to maximize net photosynthesis given the plant current conditions. Autotrophic respiration in fine roots and sapwood compartments depends on N content in these living tissue compartments. NPP is calculated yearly as gross photosynthesis minus autotrophic respiration. Annual NPP is allocated to plant's compartments growth after deduction of 10 % for reproductive costs. Leaf N demand drives total N demand as N demand for each biomass compartment is calculated to maintain a relative difference with leaf N concentration. If N availability cannot meet N demand, N uptake is equal to N availability and the labile N storage pool is used to fill the deficit. If this is still insufficient to meet N demand, the carboxylation capacity is reduced to the level that can be sustained given the current leaf N (Smith et al. 2014). More detailed information about the initial LPJ-GUESS model and the N cycle module in particular can be found in Ahlström et al. (2012) and Smith et al. (2014), respectively. LPJ-GUESS was run for the period 1900 to 2015 with an initial 500-year spin-up to get vegetation and soils into equilibrium. LPJ-GUESS was forced with monthly CRUNCEP climate (Viovy 2018), historical land-use change (Hurtt et al. 2011) and N deposition (Lamarque et al. 2013). The spatial resolution of the map was 0.5 decimal degree.

5.2.2 Random forests map

We compared the foliage N output of the two GVMs to a canopy N (%N) map in European forests based on remote sensing and environmental variables. The canopy N map was calibrated using 818 forest plots from the intensive monitoring program (level II) of the ICP Forests database (International Co-operative Program on Assessment and Monitoring of Air Pollution Effects on Forests, www.icp-forests.net). The plots were spread across European forests and were selected to represent European forests' diversity. The plots were sampled every two years between 1990-2014. The model was trained using all forest plots, i.e. deciduous broadleaf forest (DBF), evergreen needleleaf forest (ENF), evergreen broadleaf forest (EBF) and mixed plots. The canopy N map was modelled using the random forests machine learning algorithm using both remote sensing and environmental predictor variables. The environmental variables included were bioclimatic variables – WorldClim2 (Fick and Hijmans 2017), soil properties – Soilgrids250m Global Soil Information (Hengl et al. 2017), altitude (European Environment Agency 2013), N deposition (Lamarque et al. 2013) and land cover (Defourny et al. 2016). The remote sensing variables were vegetation indices and NIR from MODIS and MERIS sensors, the MOD13Q1 NDVI, EVI and NIR MODIS products (Didan 2015) as well as the MERIS-MTCI (Dash and Curran 2004). The spatial resolution of the modelled map was 300 m. More information about the methods can be found in Loozen et al. (2020). The canopy N map will hereafter be referred to as the random forests model.

5.2.3 Data preprocessing

The vegetation model simulations yielded for the O-CN model PFT-specific monthly foliage N (%N) values over the period 1980-2017 and for the LPJ-GUESS model PFT-specific yearly foliage C:N values over the period 1980-2015. The data pretreatment process was intended to make the output of the GVMs comparable with each other and comparable with the random forests model canopy N map. In this process, the obtained vegetation model values, i.e. monthly and yearly averages, were temporally aggregated in order to obtain one final foliage N (%N) for each vegetation model. This was done because the random forests model map was calibrated using long-term canopy N values.

The first step of the temporal aggregation was different between the two models. For the O-CN model, we computed yearly averages from the monthly simulated values for all forest PFTs separately at each of the grid cells. For the LPJ-GUESS yearly simulation results, we converted the C:N values into canopy N assuming a ratio of 0.48 gC/gDM, which is equal to the value used in the O-CN model:

$$\text{canopy } N = \frac{0.48}{\text{canopy } C:N} \times 100, \quad (\text{Equation 5.1})$$

Alongside the foliage N and C:N values, the output of the GVMs included fractional PFT cover (FPC) information. In O-CN, the FPC is fixed over the years, while in LPJ-GUESS model, the FPC varies yearly. For both models and for each year, we computed a weighted canopy N mean using the FPC as weight for the PFTs. In the temporal aggregation of the vegetation model outputs, we restricted the analysis to forest PFTs in order to make the results comparable with the random forests model, which was calibrated using canopy N data from forest plots exclusively. More specifically, for the O-CN model, the PFTs included in the

aggregation were boreal broadleaved summergreen (BBS), boreal needleleaved evergreen (BNE), boreal needleleaved summergreen (BNS), temperate broadleaved evergreen (TeBE), temperate broadleaved summergreen (TeBS) and temperate needleleaved evergreen (TeNE), respectively. For the LPJ-GUESS model, the PFTs were: boreal needleleaved evergreen (BNE), boreal shade-intolerant needleleaved evergreen (BINE), boreal needleleaved summergreen (BNS), temperate broadleaved evergreen (TeBE), temperate shade-intolerant broadleaved summergreen (IBS), temperate broadleaved evergreen (TeBE), respectively. Using the obtained yearly PFT weighted canopy N maps, we calculated the long-term average over the period 1980-2015 for both vegetation models.

The three canopy N maps, i.e. O-CN, LPJ-GUESS and the random forests model maps, were then resampled to the spatial resolution of the O-CN grid, i.e. 1 degree, using the bilinear interpolation method. The data pre-processing was done in the R environment, using the raster package (Hijmans 2018; R Core Team 2019).

5.3 Results

5.3.1 Spatial patterns in canopy N

The obtained canopy N maps for European forests are shown in Figure 5.1. The three maps show different but overlapping ranges of values. The canopy N from the random forests model shows a canopy N range between 1.2 and 2.2 %N, while the maps obtained from either the LPJ-GUESS simulation or the O-CN simulation show a slightly larger range of values, 0.8-2.7 %N and 1.0-2.5 %N, respectively.

More importantly, the obtained canopy N maps show contrasting spatial patterns. The canopy N map from the O-CN simulation shows relatively lower canopy N values in the central and northern part of Europe, with canopy N ranging from 1.0 %N to 1.8 %N. In Spain and the south of Italy, the canopy N values are higher, between 1.8 %N and 2.4 %N. Opposite to the canopy N map from the O-CN model, the map obtained from the LPJ-GUESS model presents relatively higher canopy N values in mid-western Europe while producing lower values in both the south and north of Europe. In central Europe, the canopy N values are between 2.0-2.8 %N, while in the north of Europe, the range of values is between 0.8 and 1.6 %N. Around the Mediterranean region, the canopy N map from LPJ-GUESS shows notably very little spatial variation, with values between 1.6 %N and 1.7 %N. Similar to the LPJ-GUESS canopy N map, the canopy N map obtained from the random forests model shows higher canopy N values, i.e. between 1.8 and 2.2 %N, in mid-western and eastern Europe, while showing lower values, i.e. between 1.2 and 1.6 %N, around the Mediterranean region and in the south of Sweden.

The contrast between the three canopy N maps becomes more clear in the difference-maps presented in Figure 5.2. Figure 5.2a shows that the values from the canopy N map produced by the random forests model are in general higher (0.3-0.5 %N) than the canopy N map from the O-CN simulation, with the exception of the Iberic Peninsula, where the values from the O-CN model are up to 1 %N higher than the random forests model. On the contrary in Figure 5.2b, the random forests map shows lower values than the LPJ-GUESS map in most of central Europe (between 0-0.8 %N lower) while showing higher values in France, Italy and the south of Sweden (0-0.2 %N higher). The random forests model also predicts notably higher values of

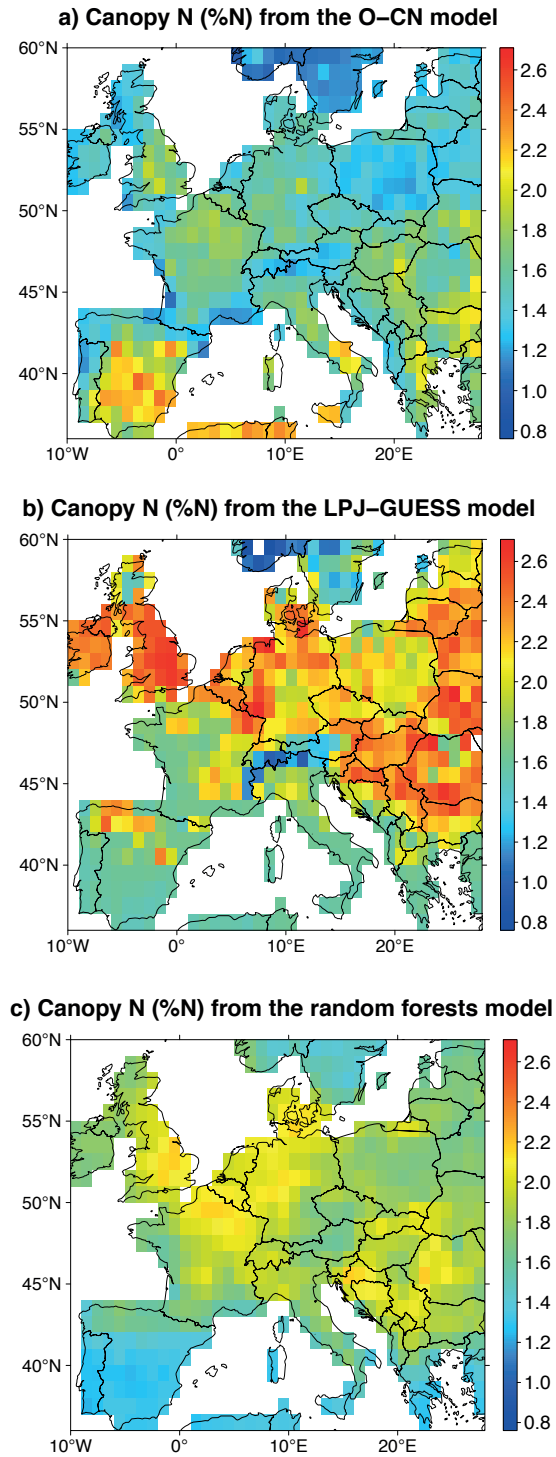


Figure 5.1. Canopy nitrogen (%N) maps in European forests calculated from a) the O-CN model, b) the LPJ-GUESS model and c) the random forests model.

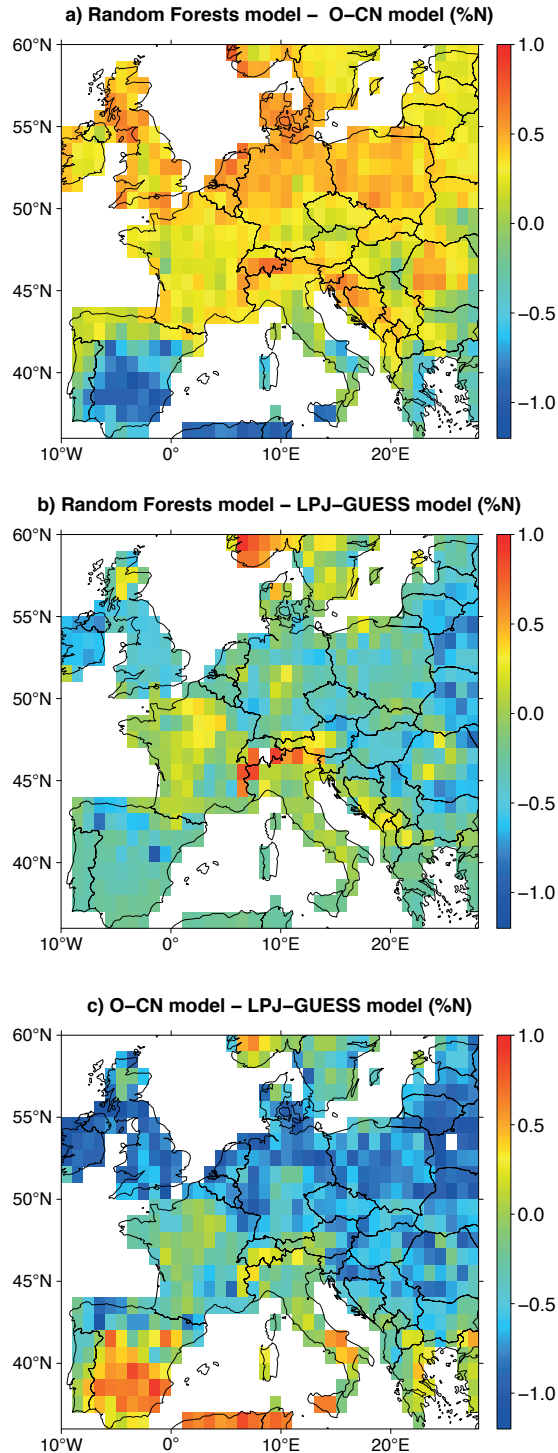


Figure 5.2. Canopy nitrogen (%N) difference maps in European forests calculated from a) the random forests model – the O-CN model, b) the random forests model – the LPJ-GUESS model and c) the O-CN model – the LPJ-GUESS model.

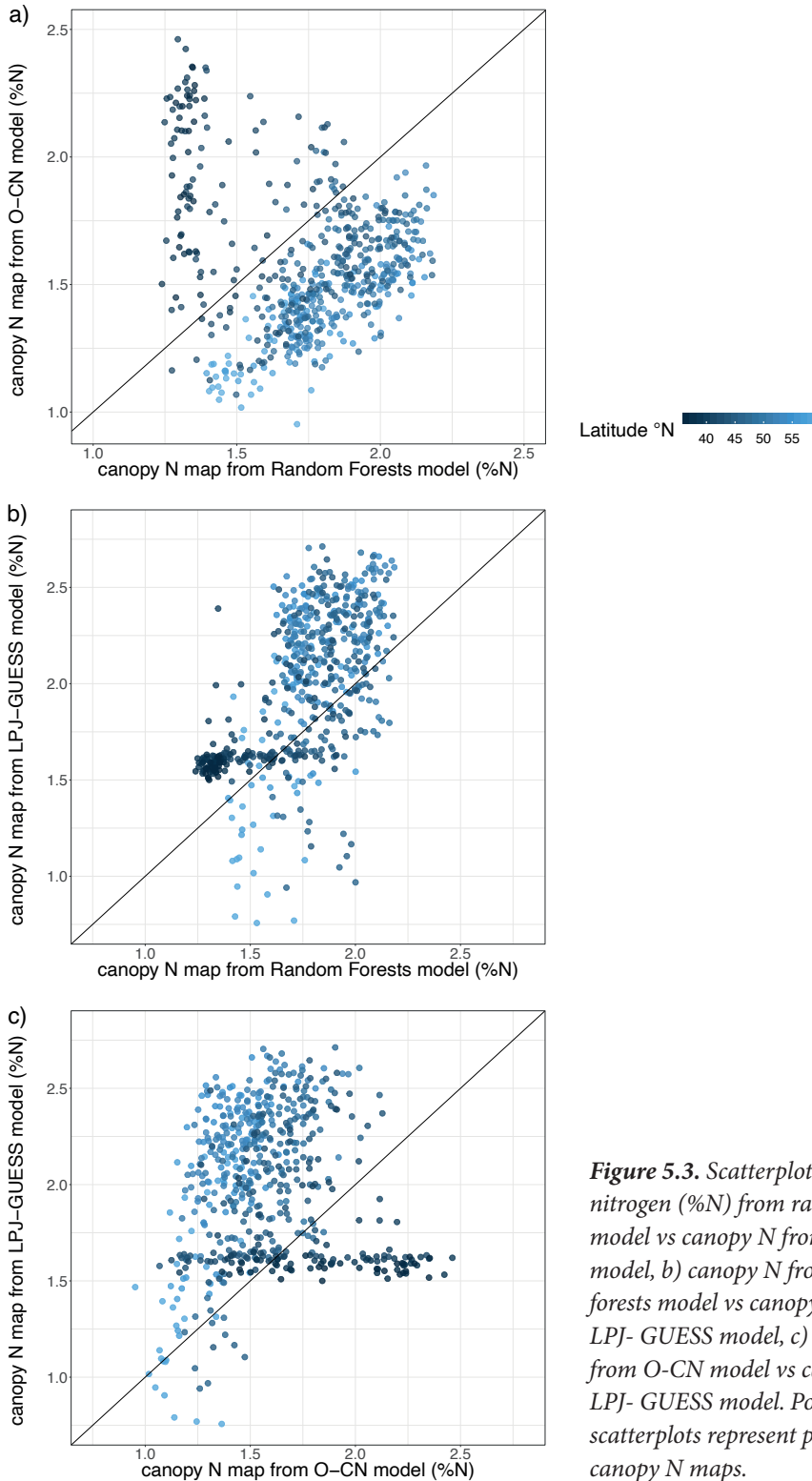


Figure 5.3. Scatterplots of a) canopy nitrogen (%N) from random forests model vs canopy N from O-CN model, b) canopy N from random forests model vs canopy N from LPJ- GUESS model, c) canopy N from O-CN model vs canopy N from LPJ- GUESS model. Points in the scatterplots represent pixels in the canopy N maps.

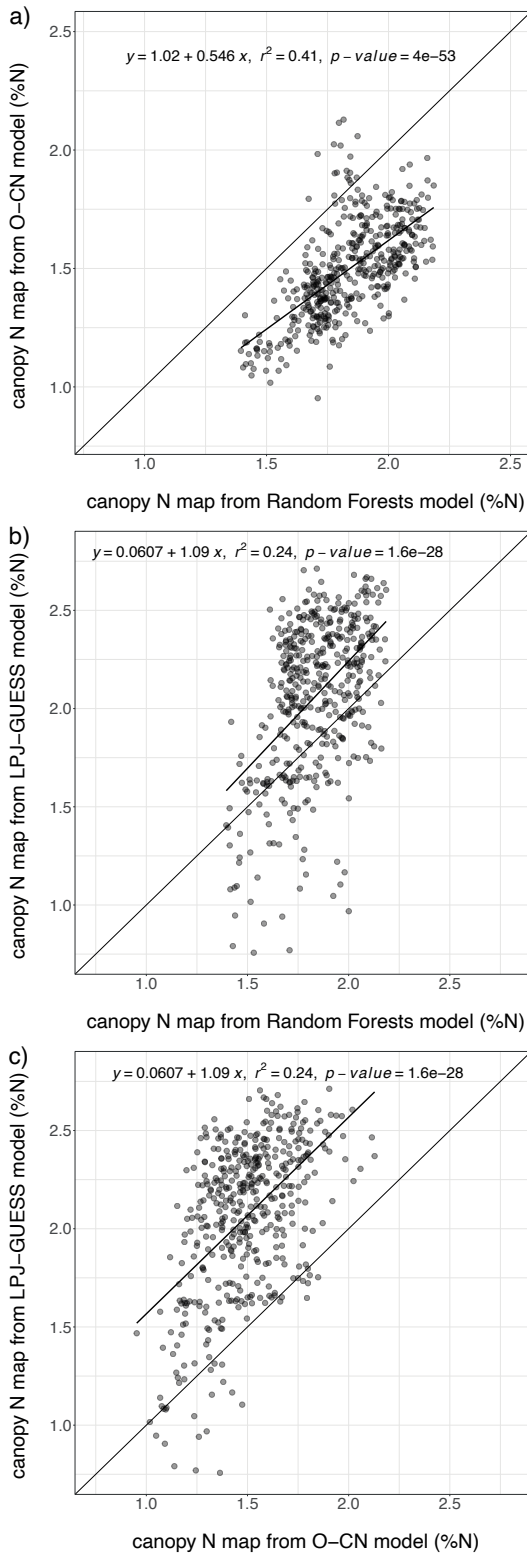


Figure 5.4. Scatterplots of a) canopy nitrogen (%N) from random forests model vs canopy N from O-CN model, b) canopy N from random forests model vs canopy N from LPJ- GUESS model, c) canopy N from O-CN model vs canopy N from LPJ- GUESS model. Points in the scatterplots represent pixels in the canopy N maps with latitude > 43 °N.

canopy N in the Alps (0.5-1 %N), compared to the LPJ-GUESS model. The difference between the canopy N predicted by the two vegetation models (Figure 5.2c) shows that the LPJ-GUESS model predicts higher values in northeastern Europe while the O-CN model predicts higher canopy N values in Spain and Italy.

5.3.2 Relationships between modelled canopy N values

Figure 5.3 presents the pixel-based scatterplot comparisons of the obtained canopy N maps. The comparison between the canopy N map from the random forests model and the canopy N map from the O-CN simulation (Figure 5.3a) shows a reasonable agreement between both maps, except for the southernmost pixels. For the pixels located in the southern region of Europe, the O-CN map shows higher values (1.5-2.5 %N) than the random forests model map (1.25-1.4 %N). Most of the other pixels are located below the identity line and show a linear trend between the values from both maps. A linear regression between the two datasets for pixels with latitude higher than 43 degrees north gives a r^2 equal to 0.41 ($p < 0.000$, Figure 5.4a). The 43 degree latitude limit was chosen by visually inspecting the O-CN canopy N map (Figure 5.1a). The region that is subsequently excluded corresponds to the Mediterranean area. Regarding the comparison between the subsequent maps produced by LPJ-GUESS and the random forests model, a dense cluster of points around the 1.6 %N value is clearly visible in the scatterplot (Figure 5.3b). These points are related to the region with little spatial variation around the Mediterranean region observed in section 5.3.1. As was done for Figure 5.4a now for Figure 5.4b, the southernmost pixels are excluded ($< 43^\circ\text{N}$) from the scatterplot between LPJ-GUESS and the random forests model. By excluding the Mediterranean region, again a clear linear relation between the two datasets emerged ($r^2 = 0.24$, $p < 0.000$). The scatterplot between the two vegetation models (Figure 5.3c) shows that canopy N values from the LPJ-GUESS simulation are in general higher than their O-CN counterparts with an exception for pixels located in southern Europe. When excluding southernmost pixels ($< 43^\circ\text{N}$), the scatterplot also shows a linear trend (Figure 5.4c, $r^2 = 0.28$, $p < 0.000$).

5.4 Discussion

5.4.1 Canopy N spatial pattern

The canopy N maps obtained from the two vegetation models, i.e. O-CN and LPJ-GUESS, showed similarities and differences compared to the random forests model map. The canopy N map obtained from the O-CN model showed similarities with the random forests map as both maps had largely overlapping canopy N value ranges, i.e. between 1.2 and 2.2 %N for the random forests map and between 1.0-2.5 %N for the O-CN map, respectively. Regarding the spatial patterns, there was in general a good agreement between the O-CN and random forests canopy N map in the northern part of Europe. The difference map (Figure 5.2a) showed that the canopy N values in the random forests map were higher than the O-CN map by a range between 0.1-0.6 %N. In the south of Europe, however, and in the Iberic peninsula especially, the differences between the two models were larger and in opposite directions. The O-CN model predicted lower values, on average between 0.5-1.2 %N lower values than the random forests model in the southern region (Figure 5.2a). In the northern region of Europe, the

results predicted by the O-CN and random forests models were in relatively good agreement. This is demonstrated by Figure 5.4a, which showed a significant linear relationship ($r^2 = 0.41$) between the canopy N values predicted by both models, once the southernmost pixels ($< 43^\circ\text{N}$) were excluded from the relationship.

The predicted map of the LPJ-GUESS model was also compared to the canopy N map produced by the random forests model. The range of canopy N values predicted by the LPJ-GUESS model was again overlapping but was also larger than the range of values produced by the random forests model; between 1.2-2.2 %N and 0.8-2.7 %N, respectively. The canopy N maps predicted by the two models showed similar spatial patterns, relatively higher canopy N values in the northern region of Europe while, in the south of Europe, the models predicted lower values in their respective ranges. The pixel-wise difference between the canopy N values predicted by two models varied between 0.5-0.5 %N in most of the region studied (Figure 5.4b). The main difference between the canopy values predicted by the random forests model and the LPJ-GUESS model was the lack of spatial variation observed in the south of Europe in the predicted LPJ-GUESS canopy N map in contrast to the random forests map. In the south of Europe, the canopy N values predicted by the random forests model showed spatial variation, between 1.4 and 1.8 %N, whereas the LPJ-GUESS model predicted values centered around 1.6 %N for the whole region. This is best illustrated in Figure 5.3b where a cluster of points is present around the 1.6 %N value. The lack of spatial variation is surprising as the rest of the study area showed apparent spatial variation and as the canopy N map from LPJ-GUESS is the one with the largest canopy N range of values compared to the two other maps. When the southernmost pixels were excluded, the relationship between the canopy N values predicted by the two models turned significant ($r^2 = 0.24$).

When compared with each other, the canopy N values predicted by the two GVMs showed overlapping ranges, between values, 0.8-2.7 %N for O-CN and 1.0-2.5 %N for LPJ-GUESS. While, in the northern region of Europe, LPJ-GUESS predicted on average higher canopy N values than O-CN, in the southern region of Europe, it was the opposite and O-CN predicted higher values than LPJ-GUESS (Figure 5.3c and Figure 5.4c). When excluding the southernmost pixels ($< 43^\circ\text{N}$), the relationship between the predicted canopy N values showed a significant linear relationship ($r^2 = 0.28$, $p < 0.000$). The linear trends observed in Figure 5.4 highlight that the canopy N values predicted by the three models are in agreement and that there is a coherence in the canopy N representations in the GVMs studied and with the random forests maps. The representation of the N cycle, while being different across the two GVMs, predicted consistent canopy N values and spatial patterns in the north of Europe.

However, the higher values in the south of Spain of O-CN and the lack of spatial variation around the Mediterranean region in LPJ-GUESS, which are the most important differences in the spatial pattern when compared with the random forests map and when compared with each other, are a reason for concern. In the southern region, the O-CN model predicted higher canopy N relative to its range, while the LPJ-GUESS and random forests model predicted the opposite, i.e. higher canopy N values, relative to their respective range of values. Therefore, we compared the results obtained in this study to an external dataset (Sardans et al. 2016) including mean PFT values in central Spain, i.e. one of the regions where the difference in spatial pattern was particularly evident. In Sardans et al. (2016), the authors grouped foliar N concentration data from several datasets (4294 Spanish forest plots), which partly overlapped

with the plots used in Loozen et al. 2020, as they included the ICP-Forests plots located in Spain (48 plots). They analyzed the mean foliar N values by PFT. In central Spain, at a latitude lower than 43 °N, the most occurring PFTs were Mediterranean gymnosperms and Mediterranean evergreen angiosperms, for which the mean foliar N concentration was 1.05 %N and 1.43 %N, respectively. Compared to the results obtained in our study, this range of values is closer to the spatial mean obtained for the random forests model (1.3 %N) than the mean obtained by the LPJ-GUESS or O-CN model, which were equal to 1.6 %N and 2.01 %N, respectively. This comparison, while being based on a single external dataset and a restricted area, is an indication that the ranges of values observed in the random forests model have a sound basis and that the values simulated by the GVMs, for the O-CN model in particular, are overestimated in central Spain.

In both GVMs studied, leaf N concentration determines the rate of photosynthesis as well as C allocation. While the spatial pattern in canopy N values across the three models presents large similarities in the northern part of Europe, the differences observed in the southern region may influence the estimated C assimilation by forest ecosystems.

When comparing the differences in the canopy N maps, it is important to take into account that the vegetation models make different assumptions about the process representation, the parameters and the inputs used. Hence, although the model results were simulated using comparable datasets, the models' output does not only depict differences in the processes represented by the models but also differences in input used to run the models.

Moreover, while we compared the output of the GVMs to the random forests map, this reference dataset is not free from inaccuracies. The r^2 and RRMSE of the random forests map were equal to 0.62 and 0.18, respectively (Loozen et al. 2020). The sources of error are linked to the data quality of the input dataset used to tune the model. While the ICP-Forests dataset used to build the random forests model follows strict guidelines regarding plot configuration, sampling and chemical analysis methods to ensure high data quality, errors can always be present. Furthermore, the model was build using 818 forest plots. While the plots were located throughout Europe, some areas suffer from poor spatial representativity in the dataset. In eastern Europe, in particular, very few plots were available. This is also the case for Great Britain (18 plots) and the south of Norway (4 plots). In Spain, however, the number of plots was higher (48 plots). Also, the spatial predictor variables used in the random forests model are prone to uncertainties which will have propagated to the model output.

5.4.2 Outlook on future research

In this study, we showed that two vegetation models showed differences in the spatial patterns and range of values of canopy N when compared to a canopy N map that was created by using intensive forest samplings and spatially continuous predictor variables, including VIs measured by earth observation. To our knowledge, a broad-scale comparison of the spatial pattern of canopy N predicted by GVMs has not been done before. One of the follow-ups of this study would be to analyze the potential causes for the dissimilarities observed, particularly in the southern region of Europe. This could be done by using new data sources and performing additional tests such as a sensitivity analysis to identify the modelled processes in the GVMs most likely to influence the canopy N values in this area. Moreover, it would be interesting

to analyze the extent of the influence of the similarities and differences of canopy N spatial pattern on the predicted C assimilation by the two GVMs.

This work focused on forests only. However, although forests represent a major C sink, other PFTs, such as grasslands, are important for C assimilation (Gourlez de la Motte et al. 2016). For this reason, it would be interesting to study the influence of including more various PFTs, and not only forests, on the results in a future comparative study. To do so, broad-scale data of foliage N would need to be available for multiple ecosystems and PFTs.

5.5 Conclusion

In this study, we compared canopy N maps from two GVMs, i.e. O-CN and LPJ-GUESS, with the canopy N maps predicted by a random forests model relying on canopy N forest sampling in European forest plots (ICP-Forest), VIs from earth observation, and spatially continuous environmental predictors. The results showed that the maps predicted overlapping ranges of canopy N values. While in the Mediterranean region, the models predicted divergent spatial patterns and range of values, there was in general a reasonable linear agreement between the canopy N maps of the three models studied regarding the spatial patterns in the northern part of Europe. This showed that, in the northern region of Europe, while the physiological processes in the two GVMs are described differently, the predicted canopy N values are consistent among each other and with the random forests model. A future perspective of this study would be to analyze further the causes, at the level of the processes represented in the GVMs, of the differences observed. This would be particularly useful in the southern region of Europe.

5.6 References

- Ahlström, A., Schurgers, G., Arneeth, A., & Smith, B. (2012). Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections. *Environmental Research Letters*, 7, 044008. 10.1088/1748-9326/7/4/044008
- Ahlström, A., Schurgers, G., & Smith, B. (2017). The large influence of climate model bias on terrestrial carbon cycle simulations. *Environmental Research Letters*, 12, 014004. 10.1088/1748-9326/12/1/014004
- Bindoff, N.L., Stott, P.A., AchutaRao, K.M., Allen, M.R., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S., Mokhov, I.I., Overland, J., Perlwitz, J., Sebbari, R., & Zhang, X. (2013). Detection and Attribution of Climate Change: from Global to Regional. In T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P.M.e. Midgley (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press
- Chapin, F.S.I. (1987). Adaptations and physiological responses of wild plants to nutrient stress. In W.H. Gabelman, & B.C. Loughman (Eds.), *Genetic Aspects of Plant Mineral Nutrition. Developments in Plant and Soil Sciences* Dordrecht: Springer. https://doi.org/10.1007/978-94-009-3581-5_2.
- Dash, J., & Curran, P.J. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25, 5403-5413. doi:10.1080/0143116042000274015
- Defourny, P., Kirches, G., Brockmann, C., Boettcher, M., Peters, M., Bontemps, S., Lamarche, C., Schlerf, M., & Santoro, M. (2016). *Land Cover CCI Product User Guide Version 2.5* (p. 91). Louvain-La-Neuve, Belgium: UCL-Geomatics. Accessed 30 May 2019. <http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-PUG-v2.5.pdf>
- Didan, K. (2015). *MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006* [Dataset]. USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota: NASA EOSDIS Land Processes DAAC. Accessed 28 January 2019. <https://lpdaac.usgs.gov>. <https://doi.org/10.5067/MODIS/MOD13Q1.006>
- Dlugokencky, E., & Tans, P. (2018). *Trends in atmospheric carbon dioxide, National Oceanic & Atmospheric Administration* [Dataset]. Earth System Research Laboratory (NOAA/ESRL). Accessed 4 September 2018. <http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html>
- European Environment Agency (2013). *Digital Elevation Model over Europe (EU-DEM)* [Dataset]. European Environment Agency. Accessed 20 April 2015. <https://www.eea.europa.eu/data-and-maps/data/eu-dem>
- Exbrayat, J.F., Bloom, A.A., Carvalhais, N., Fischer, R., Huth, A., MacBean, N., & Williams, M. (2019). Understanding the Land Carbon Cycle with Space Data: Current Status and Prospects. *Surveys in Geophysics*, 40, 735-755. 10.1007/s10712-019-09506-2
- Fernández-Martínez, M., Vicca, S., Janssens, I.A., Sardans, J., Luysaert, S., Campioli, M., Chapin Iii, F.S., Ciais, P., Malhi, Y., Obersteiner, M., Papale, D., Piao, S.L., Reichstein, M., Rodà, F., & Peñuelas, J. (2014). Nutrient availability as the key regulator of global forest carbon balance. *Nature Climate Change*, 4, 471-476. 10.1038/nclimate2177
- Fick, S.E., & Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302-4315. doi:10.1002/joc.5086

- Fleischer, K., Wårlind, D., van der Molen, M.K., Rebel, K.T., Arneth, A., Erisman, J.W., Wassen, M.J., Smith, B., Gough, C.M., Margolis, H.A., Cescatti, A., Montagnani, L., Arain, A., & Dolman, A.J. (2015). Low historical nitrogen deposition effect on carbon sequestration in the boreal zone. *Journal of Geophysical Research: Biogeosciences*, *120*, 2542-2561. 10.1002/2015JG002988
- Friedlingstein, P., Jones, M.W., O'Sullivan, M., Andrew, R.M., Hauck, J., Peters, G.P., Peters, W., Pon-gratz, J., Sitch, S., Le Quéré, C., Bakker, D.C.E., Canadell, J.G., Ciais, P., Jackson, R.B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L.P., Currie, K.I., Feely, R.A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D.S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R.A., Hurtt, G., Ilyina, T., Jain, A.K., Joetzjer, E., Kaplan, J.O., Kato, E., Klein Goldewijk, K., Korsbakken, J.I., Landschützer, P., Lauvset, S.K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P.C., Melton, J.R., Metzl, N., Munro, D.R., Nabel, J.E.M.S., Nakaoka, S.I., Neill, C., Omar, A.M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P.P., Tian, H., Tilbrook, B., Tubiello, F.N., van der Werf, G.R., Wiltshire, A.J., & Zaehle, S. (2019). Global Carbon Budget 2019. *Earth Syst. Sci. Data*, *11*, 1783-1838. 10.5194/essd-11-1783-2019
- Gourlez de la Motte, L., Jérôme, E., Mamadou, O., Beckers, Y., Bodson, B., Heinesch, B., & Aubinet, M. (2016). Carbon balance of an intensively grazed permanent grassland in southern Belgium. *Agricultural and Forest Meteorology*, *228-229*, 370-383. <https://doi.org/10.1016/j.agrformet.2016.06.009>
- Hansen, P.M., & Schjoerring, J.K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, *86*, 542-553. doi:10.1016/S0034-4257(03)00131-7
- Harris, I., Jones, P.D., Osborn, T.J., & Lister, D.H. (2014). Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, *34*, 623-642. 10.1002/joc.3711
- Heggin, M., Kinnison, D., & Lamarque, J.-F. (2016). *CCMI nitrogen surface fluxes in support of CMIP6 – version 2.0* [Dataset]. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/input4MIPs.1125>
- Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangquan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, *12*, e0169748. 10.1371/journal.pone.0169748
- Hijmans, R.J. (2018). *raster: Geographic Data Analysis and Modeling*. R package version 2.8-4. Accessed 26 April 2019. <https://CRAN.R-project.org/package=raster>
- Huber, S., Kneubühler, M., Psomas, A., Itten, K., & Zimmermann, N.E. (2008). Estimating foliar biochemistry from hyperspectral data in mixed forest canopy. *Forest Ecology and Management*, *256*, 491-501. doi:10.1016/j.foreco.2008.05.011
- Huntzinger, D.N., Michalak, A.M., Schwalm, C., Ciais, P., King, A.W., Fang, Y., Schaefer, K., Wei, Y., Cook, R.B., Fisher, J.B., Hayes, D., Huang, M., Ito, A., Jain, A.K., Lei, H., Lu, C., Maignan, F., Mao, J., Parazoo, N., Peng, S., Poulter, B., Ricciuto, D., Shi, X., Tian, H., Wang, W., Zeng, N., & Zhao, F. (2017). Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Scientific Reports*, *7*. 10.1038/s41598-017-03818-2
- Hurtt, G.C., Chini, L.P., Froking, S., Betts, R.A., Feddema, J., Fischer, G., Fisk, J.P., Hibbard, K., Houghton, R.A., Janetos, A., Jones, C.D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D.P., & Wang, Y.P.

- (2011). Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, 109, 117. 10.1007/s10584-011-0153-2
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., & Takahashi, K. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, 93, 5-48. 10.2151/jmsj.2015-001
- Lamarque, J.F., Dentener, F., McConnell, J., Ro, C.U., Shaw, M., Vet, R., Bergmann, D., Cameron-Smith, P., Dalsoren, S., Doherty, R., Faluvegi, G., Ghan, S.J., Josse, B., Lee, Y.H., MacKenzie, I.A., Plummer, D., Shindell, D.T., Skeie, R.B., Stevenson, D.S., Strode, S., Zeng, G., Curran, M., Dahl-Jensen, D., Das, S., Fritzsche, D., & Nolan, M. (2013). Multi-model mean nitrogen and sulfur deposition from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): evaluation of historical and projected future changes. *Atmos. Chem. Phys.*, 13, 7997-8018. 10.5194/acp-13-7997-2013
- LeBauer, D.S., & Treseder, K.K. (2008). Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, 89, 371-379. 10.1890/06-2057.1
- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., & Martin, M.E. (2016). Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sensing of Environment*, 173, 174-186. doi:10.1016/j.rse.2015.11.028
- Loozen, Y., Karssenber, D., de Jong, S.M., Wang, S., van Dijk, J., Wassen, M.J., & Rebel, K.T. (2019). Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses. *International Journal of Applied Earth Observation and Geoinformation*, 75, 1-14. https://doi.org/10.1016/j.jag.2018.08.012
- Loozen, Y., Rebel, K.T., de Jong, S.M., Lu, M., Ollinger, S.V., Wassen, M.J., & Karssenber, D. (2020). Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method. *Remote Sensing of Environment*, 247, 111933. https://doi.org/10.1016/j.rse.2020.111933
- Loozen, Y., Rebel, K.T., Karssenber, D., Wassen, M.J., Sardans, J., Peñuelas, J., & de Jong, S.M. (2018). Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index. *Biogeosciences*, 15, 2723-2742. 10.5194/bg-15-2723-2018
- Lovenduski, N.S., & Bonan, G.B. (2017). Reducing uncertainty in projections of terrestrial carbon uptake. *Environmental Research Letters*, 12. 10.1088/1748-9326/aa66b8
- Majeke, B., Van Aardt, J.A.N., & Cho, M.A. (2008). Imaging spectroscopy of foliar biochemistry in forestry environments. *Southern Forests*, 70, 275-285. 10.2989/SF.2008.70.3.11.672
- McNeil, B.E., Read, J.M., & Driscoll, C.T. (2007). Foliar nitrogen responses to elevated atmospheric nitrogen deposition in nine temperate forest canopy species. *Environmental Science and Technology*, 41, 5191-5197. doi: 10.1021/es062901z
- Meyerholt, J., Zaehle, S., & Smith, M.J. (2016). Variability of projected terrestrial biosphere responses to elevated levels of atmospheric CO₂ due to uncertainty in biological nitrogen fixation. *Biogeosciences*, 13, 1491-1518. 10.5194/bg-13-1491-2016
- Moreno-Martínez, Á., Camps-Valls, G., Kattge, J., Robinson, N., Reichstein, M., van Bodegom, P., Kramer, K., Cornelissen, J.H.C., Reich, P., Bahn, M., Niinemets, Ü., Peñuelas, J., Craine, J.M., Cerabolini, B.E.L., Minden, V., Laughlin, D.C., Sack, L., Allred, B., Baraloto, C., Byun, C., Soudzilovskaia, N.A., & Running, S.W. (2018). A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sensing of Environment*, 218, 69-88. https://doi.org/10.1016/j.rse.2018.09.006

- R Core Team (2019). *R: A language and environment for statistical computing* [Computer program]. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., & Heitkönig, I.M.A. (2012). Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor. *International Journal of Applied Earth Observation and Geoinformation*, 19, 151-162. doi:10.1016/j.jag.2012.05.009
- Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Volin, J.C., & Bowman, W.D. (1999). Generality of leaf trait relationships: A test across six biomes. *Ecology*, 80, 1955-1969. doi:10.2307/176671
- Reichstein, M., & Carvalhais, N. (2019). Aspects of Forest Biomass in the Earth System: Its Role and Major Unknowns. *Surveys in Geophysics*, 40, 693-707. 10.1007/s10712-019-09551-x
- Sardans, J., Alonso, R., Carnicer, J., Fernández-Martínez, M., Vivanco, M.G., & Peñuelas, J. (2016). Factors influencing the foliar elemental composition and stoichiometry in forest trees in Spain. *Perspectives in Plant Ecology, Evolution and Systematics*, 18, 52-69. 10.1016/j.ppees.2016.01.001
- Schlerf, M., Atzberger, C., Hill, J., Buddenbaum, H., Werner, W., & Schüller, G. (2010). Retrieval of chlorophyll and nitrogen in Norway spruce (*Picea abies* L. Karst.) using imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, 12, 17-26. doi:10.1016/j.jag.2009.08.006
- Schurgers, G., Ahlström, A., Arneth, A., Pugh, T.A.M., & Smith, B. (2018). Climate Sensitivity Controls Uncertainty in Future Terrestrial Carbon Sink. *Geophysical Research Letters*, 45, 4329-4336. 10.1029/2018GL077528
- Smith, B., Wärlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, 11, 2027-2054. doi:10.5194/bg-11-2027-2014
- Viovy, N. (2018). *CRUNCEP Version 7 - Atmospheric Forcing Data for the Community Land Model* [Dataset]. Boulder, CO: Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <http://rda.ucar.edu/datasets/ds314.3/>
- Vitousek, P.M., & Howarth, R.W. (1991). Nitrogen limitation on land and in the sea: How can it occur? *Biogeochemistry*, 13, 87-115. 10.1007/BF00002772
- Wärlind, D., Smith, B., Hickler, T., & Arneth, A. (2014). Nitrogen feedbacks increase future terrestrial ecosystem carbon uptake in an individual-based dynamic vegetation model. *Biogeosciences*, 11, 6131-6146. 10.5194/bg-11-6131-2014
- Wieder, W.R., Cleveland, C.C., Smith, W.K., & Todd-Brown, K. (2015). Future productivity and carbon storage limited by terrestrial nutrient availability. *Nature Geoscience*, 8, 441. 10.1038/ngeo2413
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., & Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, 428, 821-827. doi:10.1038/nature02403
- Xue, L., Cao, W., Luo, W., Dai, T., & Zhu, Y. (2004). Monitoring Leaf Nitrogen Status in Rice with Canopy Spectral Reflectance. *Agronomy Journal*, 96, 135-142. 10.2134/agronj2004.0135
- Zaehle, S. (2013). Terrestrial nitrogen-carbon cycle interactions at the global scale. *Philosophical Transactions of the Royal Society B*, 368. <http://doi.org/10.1098/rstb.2013.0125>

- Zaehle, S., & Dalmonech, D. (2011). Carbon-nitrogen interactions on land at global scales: Current understanding in modelling climate biosphere feedbacks. *Current Opinion in Environmental Sustainability*, 3, 311-320. 10.1016/j.cosust.2011.08.008
- Zaehle, S., Friedlingstein, P., & Friend, A.D. (2010). Terrestrial nitrogen feedbacks may accelerate future climate change. *Geophysical Research Letters*, 37. <https://doi.org/10.1029/2009GL041345>
- Zaehle, S., & Friend, A.D. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, 24. 10.1029/2009GB003521
- Zaehle, S., Medlyn, B.E., De Kauwe, M.G., Walker, A.P., Dietze, M.C., Hickler, T., Luo, Y., Wang, Y.-P., El-Masri, B., Thornton, P., Jain, A., Wang, S., Warlind, D., Weng, E., Parton, W., Iversen, C.M., Gallet-Budynek, A., McCarthy, H., Finzi, A., Hanson, P.J., Prentice, I.C., Oren, R., & Norby, R.J. (2014). Evaluation of 11 terrestrial carbon-nitrogen cycle models against observations from two temperate Free-Air CO₂ Enrichment studies. *New Phytologist*, 202, 803-822. 10.1111/nph.12697

Chapter 6

Synthesis

6.1 Context

Nitrogen (N) is an essential and limiting nutrient for plant growth (LeBauer and Treseder 2008; Vitousek and Howarth 1991). N cycle and N availability influences the C cycle and C assimilation by the terrestrial biosphere (Ciais et al. 2013; Fernández-Martínez et al. 2014; Wieder et al. 2015). In order to quantify this influence, data on the N cycle is needed. In this context, this thesis addressed the scientific problem of the gap in spatially explicit information on canopy N at large scale.

The studies presented in this thesis explored the estimation of canopy N with remotely sensed VIs, as well as environmental variables, across vegetation types and spatial scales. Canopy N estimated with VIs and environmental variables were also compared with canopy N simulated by global vegetation models (GVMs). The research questions addressed in this thesis and presented in the introduction chapter were:

How well can we estimate canopy nitrogen (N) across spatial scales using vegetation indices (VIs) from remote sensing and environmental variables?

- (i) What is the accuracy of canopy N estimated from remotely sensed vegetation indices (VIs)? (Chapter 2, 3 and 4)
- (ii) How will canopy N estimation from remote sensing be influenced if environmental variables are included as predictors? (Chapter 4)
- (iii) How does canopy N estimated from remote sensing and environmental variables compare with foliage nitrogen simulated by global vegetation models (GVMs)? (Chapter 5)

The following sections address each sub-question separately. Future perspectives are then also discussed.

6.2 Remotely sensed vegetation indices for canopy N estimation

- (i) **What is the accuracy of canopy N estimated from remotely sensed Vegetation Indices (VIs)?**

Remotely sensed vegetation indices (VIs) have been used to estimate canopy N (%N) in a variety of plant species and ecosystems. This can be attributed to their ease of use as they are calculated as a ratio of reflectance bands. As such, they can be computed from a range of different sensors, with either narrow or broadband detection. The sensor can either be handheld or onboard a plane or satellite.

In this thesis, we explored and assessed the accuracy of VIs for canopy N estimation in different conditions, spanning a range of scales and vegetation types, from a controlled laboratory experiment, to regional and continental studies with increasing levels of heterogeneity regarding the vegetation studied. Chapter 2 describes a laboratory experiment during which the reflectance spectra of a grass species, *Holcus lanatus*, was measured under controlled conditions. Using the measured reflectance, 60 existing VIs were obtained, as well as VIs specifically developed for this experiment. The reflectance spectra were also resampled to

the band settings of six satellite's sensors to assess the accuracy of canopy N estimation when VIs were calculated with broader band settings. Canopy N and VIs were related using linear or log-transformed regressions.

In chapter 3, we estimated canopy N and canopy N content (g m^{-2}) in a Mediterranean forest ecosystem at regional scale using a red-edge based VI, the MERIS Terrestrial Chlorophyll Index (MTCI). The MTCI was obtained from the MERIS sensor aboard ESA-Envisat satellite. The study was set in Catalonia and exploited canopy N from 841 forest plots measured by CREAM during 1988-2001.

In chapter 4, we estimated canopy N in European forests at continental scale using either only satellite-based variables or both VIs and environmental variables as predictors. While in chapter 2 and 3, canopy N was estimated using linear regression, in chapter 4, canopy N was estimated using the random forests algorithm.

In estimating canopy N with VIs, it is important to assess how accurate the estimation is. To do this, across the three chapters, coefficients of determination (r^2) were presented to evaluate the results. Below, these measures are discussed and compared across the three studies. The obtained r^2 are summarized in Table 6.1. Although the tabulated values are informative, care should be taken comparing r^2 values between the studies in the different chapters as the spatial and temporal resolutions are different as well as the spatial extent of the data over which the coefficients are calculated. This will be further discussed below.

The results of chapter 2 showed that among the 60 existing VIs tested to estimate canopy N, 32 VIs showed a significant relationship with canopy N, with r^2 between 0.16-0.44. The VIs optimized for the experiment showed higher r^2 , between 0.58-0.69, for the narrow band sensor. When resampling the reflectance spectra to the band settings of the satellite sensors, the r^2 of the linear regressions between canopy N and the VIs ranged between 0.21-0.58 for the existing VIs, and between 0.47-0.67 for the optimized VIs. These results showed that the optimized VIs showed in general higher r^2 than existing VIs. The range of r^2 values obtained were in agreement with the results observed in other studies (Pacheco-Labrador et al. 2014; Tian et al. 2011). Among the existing VIs tested, several VIs based on the red-edge region of the reflectance spectrum showed significant relationship with canopy N. The r^2 of the relationships between the red-edge VIs and canopy N ranged between 0.26-0.33. In particular, the MTCI was related to canopy N with a r^2 equal to 0.31.

In chapter 3, the relationship between MTCI at 1 km spatial resolution and canopy N across PFTs showed an r^2 of 0.32, while the r^2 of the relationship with canopy N content was lower, equal to 0.17. These results are similar or higher than published studies estimating canopy N with MTCI in grasslands and temperate forests (Ramoelo et al. 2012; Wang et al. 2016). The observed r^2 is also similar to the r^2 obtained in chapter 2 for canopy N estimation with MTCI. Although the scale of measurement is different, this indicates consistency in the results presented. Chapter 3 also highlighted the difference between PFTs regarding canopy N estimation with MTCI. The relationship between MTCI and canopy N was stronger for DBF plots ($r^2 = 0.24$) than for ENF plots ($r^2 = 0.10$), while the contrary was true for the relationship between MTCI and canopy N content ($r^2 = 0.19$ for ENF; insignificant for DBF).

In chapter 4, canopy N estimated across PFTs with remotely sensed variables showed an r^2 equal to 0.60. This result was comparable to those observed in experiments using similar

Table 6.1. Summary of the coefficients of determination (r^2) obtained in the different studies presented. PFT = plant functional type, DBF = deciduous broadleaf forest, ENF = evergreen needleleaf forest, VI = Vegetation index, MTCI = MERIS Terrestrial Chlorophyll Index, EVI = Enhanced Vegetation Index, NDVI = Normalized Difference Vegetation Index.

Chapter	Scale	Vegetation type	VI tested	r^2
Chapter 2	Laboratory experiment	Grass	Existing VIs – narrow band sensor	0.16-0.44
	Laboratory experiment	Grass	Optimized VIs – narrow band sensor	0.58-0.69
	Laboratory experiment	Grass	Existing VIs – resampled spectra	0.21-0.58
	Laboratory experiment	Grass	Optimized VIs – resampled spectra	0.47-0.67
Chapter 3	Regional	All PFTs	MTCI	0.32
	Regional	DBF	MTCI	0.24
	Regional	ENF	MTCI	0.10
Chapter 4	Continental	All PFTs	MTCI, EVI, NDVI	0.60
	Continental	DBF	MTCI, EVI, NDVI	0.09
	Continental	ENF	MTCI, EVI, NDVI	0.44

machine learning methods to estimate canopy N across large areas (Moreno-Martínez et al. 2018). The results within PFTs showed differences with what was observed in chapter 3. The r^2 observed for ENF plots was equal to 0.44, which was higher than the r^2 observed for DBF plots ($r^2 = 0.09$). In chapter 3, the opposite was observed as the relationship between MTCI and canopy N for DBF plots was stronger than its counterpart for ENF plots.

The results showed that the accuracy, in terms of r^2 , of canopy N estimation across the three studies was in a large range, between $r^2 = 0.09$ -0.69. As noted above, when comparing the accuracy of the relationships across the three studies, we should however keep in mind that they are not completely comparable as the results were obtained under different conditions, sensors, scales and settings. While, in chapter 2, the reflectance measurements were taken simultaneously with the foliar samples, the larger spatial areas covered in chapter 3 and 4 implied that all the forest plots could not be sampled at once and simultaneously with one satellite sensor overflight. In these studies, both the canopy N and VIs values were averaged temporally, which represents a major difference in methodology compared to chapter 2. The difference in scale also appeared to influence the relationships observed. Chapter 2 used VIs and canopy N values representative for a support size of 12 cm², while in chapter 3 and 4, values referred to pixels of 1 km² and 0.09 km², respectively. These datasets refer to spatial variations in the canopy at different scales. Both these differences in methodology and spatial representativity affect the accuracy of the relationships observed. The type of model also influences the accuracy. For example, in chapter 4, the r^2 observed for all the PFTs is higher than in chapter 3. However, in chapter 4, the model is based on more explanatory variables, which are expected to contain more information than a single explanatory variable.

One of the remaining uncertainties related to canopy N remote sensing involves the precise mechanism explaining the observed relationships. This has led to controversy over the cause behind the observed relationships. In forest environments, in particular, the study by Ollinger et al. (2008) investigating the remote detection of canopy N in boreal forests, was criticized in several subsequent publications (Knyazikhin et al. 2013a; Knyazikhin et al. 2013b;

Knyazikhin et al. 2013c; Townsend et al. 2013). The critiques were based on the ground that the observed relationships between canopy N and NIR albedo were solely the consequences of canopy structural effects resulting from the presence of a combination of several PFTs in the forests studied. The results presented in this thesis provide new and valuable information in this scientific debate. The database with canopy N data from forest plots used in chapter 3 and chapter 4 mainly included monospecific plots. This made it possible to study the relationships between canopy N and VIs for single PFTs and even single species, for a restricted number of plots in chapter 3. The result showed that significant relationships between canopy N and VIs could be defined for single PFTs. This was the case for DBF and ENF plots in Mediterranean forests in chapter 3 as well as for ENF plots in European forests in chapter 4. While these results do not provide further information about the actual mechanisms behind canopy N remote sensing, the existence of significant relationships between VIs and canopy N for PFTs taken separately contradicts, in these specific conditions, the idea suggested before that only structural effects are driving canopy N estimation with remote sensing in forest environments. Uncertainties remain, however, about the reason why the observed accuracy for one PFT differs, and sometimes greatly in the case of DBF plots, between chapter 3 and chapter 4.

Similarly, in chapter 2, the results showed that there was a difference in accuracy between the optimized and existing VIs for canopy N estimation. Moreover, 28 out of the 60 existing VIs tested did not show a significant relationship with canopy N, while being included in the analysis because these VIs were successful at estimating canopy N in previous studies. This can also be observed in other studies at local scale comparing existing and previously well performing VIs. Among the multiple VIs tested for canopy N estimation, several show poor accuracy when tested in different conditions than the ones they were developed for (Pacheco-Labrador et al. 2014; Tian et al. 2011). The changes in conditions include difference in plant species, spatial resolution, sensor or reflectance band settings. This inconsistency in the accuracy of VIs across studies may be a sign that the models developed are dependent on specific conditions. This raises questions regarding the generalization of the relationships observed. Understanding the causes of the variation in accuracy is essential for future real-world applications of canopy N estimation with VIs.

6.3 Including environmental variables as predictors

(ii) How will canopy N estimation from remote sensing be influenced if environmental variables are included as predictors?

In chapter 4, we estimated canopy N at continental scale in European forests using either only remote sensing or both remote sensing and environmental variables as predictors. Including environmental variables alongside remote sensing variables to estimate canopy N has already been done in previous studies, especially in Savannah ecosystems (McNeil et al. 2012; Ramoelo et al. 2011; Ramoelo et al. 2012; Ramoelo et al. 2013).

Integrating environmental variables as predictors in the predictive models is motivated by the heterogeneity of the ecosystems studied (Ramoelo et al. 2011). This method is thus suited for estimating canopy N across a large spatial area, such as European forests which encompass a wide variety of climate, ecosystems and edaphic conditions. Recently, predictive

models based on both remote sensing and climate variables successfully estimated canopy N at global scale (Moreno-Martínez et al. 2018). Global data-driven studies have shown that foliar N concentration covary with climate factors including temperature (Reich and Oleksyn 2004). In forest environments, other environmental variables have been shown to be correlated with foliar N, among which soil properties and N deposition (Sardans et al. 2016b; Sardans et al. 2015). In chapter 4, the environmental variables included were climate, land cover, N deposition, altitude and soil properties.

In chapter 4, we analyzed the influence on the accuracy of canopy N estimation by adding environmental variables. In the model estimating canopy N for all PFTs combined, the r^2 of the random forests model was not strongly influenced by including environmental variables in the models ($r^2 = 0.62$) compared with the models with remote sensing variables only ($r^2 = 0.60$). The influence was stronger on the accuracy of the predictive models for PFTs taken separately, which was especially the case for DBF plots. For ENF forest plots, the r^2 increased from 0.44 to 0.45 when including environmental variables. For DBF, the influence of including environmental variables in the predictive model was strongest. While the r^2 was equal to 0.09 with remote sensing only predictors, it increased to 0.39 with the inclusion of environmental variables. These results showed that including environmental variables was more beneficial for the DBF model compared to the ENF or all plots models. In any case, including environmental variables had a positive influence on the accuracy of the models. Similar results were also found in studies evaluating the influence of adding environmental variables on the predictive models. For example, in a study on Savannah grasses at regional scale, including altitude in the predictors increased the accuracy of the predictive model compared to a model based on a red-edge VI only (Ramoelo et al. 2012).

An outcome highlighted in chapter 4 is thus that the inclusion of environmental variables as predictors in the random forests model improved the accuracy of canopy N prediction. Following this, would it be possible to map canopy N at large scale using environmental variables only? In the literature, mapping foliar traits using exclusively environmental variables as predictors has already been done at global scale (Butler et al. 2017). Remote sensing variables are nevertheless relevant for canopy N mapping. While environmental variables data can be spatially interpolated from local measurements, e.g. climate data (Fick and Hijmans 2017) and soil data (Hengl et al. 2017), spectral reflectance from satellite sensors are directly measured at every point on a continuous grid, which provides information about the vegetation at a finer scale, depending on the spatial resolution of the satellite. Moreover, compared to environmental data obtained from one-time or infrequent measurements, spectral reflectance from satellite sensors are measured on a regular basis. Environmental variables and remote sensing variables should thus not be opposed as these two sources of information are complementary.

The research presented in this thesis is grounded in the broader scientific ambition to estimate canopy N from remote sensing. This ambition revolves around the fact that data on the N cycle is lacking at large scale. In this context, the results presented in chapter 4 contribute to solving this scientific problem as they showed that it was possible to map canopy N at continental scale with relatively high accuracy ($r^2 = 0.62$). This result may be leading the path towards mapping canopy N at global scale.

The methods used in chapter 4 could be applied to map canopy N at global scale. To do so, while data availability is of foremost importance, the quality of the datasets used is equally important and will influence the models developed. The results of chapter 4 showed that among the environmental variables included in the models, the ones with the most influence on the random forests models were bioclimatic, soil and N deposition variables. Improving the quality of these predictor variables has thus the highest potential of having a positive influence on the quality of the models developed. One of the aspects of the quality of the variables that could be improved is the spatial resolution. For the N deposition variable especially, for which the spatial resolution of the product used in chapter 4 was equal to 50 km, it would be interesting to test the influence of using a product with a higher spatial resolution, if available. As the N deposition variables were among the most influential for the models of DBF vegetation, using an N deposition map with a higher spatial resolution is recommended and may improve the accuracy of these models, which showed the lowest accuracy compared to the other PFTs.

6.4 Remote sensing canopy N estimates for global vegetation models

(iii) How does canopy N estimated from remote sensing and environmental variables compare with foliage nitrogen simulated by global vegetation models?

For calibrating and evaluating GVMs, data are needed but, although ample data on the C cycle are available for model evaluation, data relative to the N cycle are lacking at global scale. Thus far, limited evaluation of the N modules of GVMs has been done, mainly by validation against their ability to reproduce observed trends in the C cycle (Zaehle and Dalmonech 2011) but so far no attempts were carried out to evaluate and compare how different GVMs perform in predicting canopy N spatial patterns.

In this context, data from remote sensing have been identified as a promising source to provide information about essential environmental variables (Giuliani et al. 2020) and data from earth observation holds the potential to provide spatially explicit information on the N cycle at large scale (Lepine et al. 2016; Moreno-Martínez et al. 2018; Ollinger et al. 2008). Earth observation data have already been used to improve GVM predictions related to the C cycle (Exbrayat et al. 2019; Scholze et al. 2017). For example, remote sensing based estimation of gross and NPP have been compared to model simulations (Exbrayat et al. 2019).

For the O-CN model, the predicted foliage N values have also already been evaluated against site-scale data in the temperate region (12 sites). The results showed that there was an agreement between the simulated and observed foliage N values for both broadleaved and needleleaved PFTs (Zaehle and Friend 2010). The LPJ-GUESS model was also evaluated against site-scale data (28 sites). Although the model was able to capture the difference between the PFTs, the difference in foliage N within PFTs was not well simulated by the model (Fleischer et al. 2015). For both of these GVMs, simulated canopy N was compared with the spatial pattern of N estimated from remote sensing.

Chapter 5 presented this comparison between a canopy N map based on remote sensing and environmental variables and canopy N simulated by O-CN and LPJ-GUESS. The canopy N maps were compared based on the predicted range of canopy N values as well as the predicted spatial pattern at European scale. The results highlighted the large similarities observed

regarding spatial pattern and overlapping range of values in the northern region of Europe. In the Mediterranean region, contrasting spatial patterns were found across the three models compared. We also compared the range of values obtained in central Spain to an external dataset (Sardans et al. 2016a). The comparison showed that the range of values observed in the canopy N map based on remote sensing and environmental variables were closer to the average values from the external dataset than the values obtained from either the LPJ-GUESS or O-CN model. While this observed difference in the three canopy N maps is a source of concern, the cause behind this discrepancy is still unknown. Further research into the modelled processes and input data used is needed to untangle the reasons behind this divergence in the canopy N values in southern Europe.

Including the N cycle is a recent addition in GVMs and there are still uncertainties in the way N cycle related physiological processes should be represented in the vegetation models (Davies-Barnard et al. 2020). The study presented in chapter 5 evaluated the canopy N predicted by two GVMs. Following this, and provided the quality of the product is good enough, the availability of remote sensing based data, such as canopy N maps, provide the opportunity to improve the model representation of the processes that influence the simulated canopy N values.

6.5 Future perspectives

The studies presented in this thesis addressed the estimation of canopy N at multiple scales, from a laboratory experiment in chapter 2, to regional and continental estimation of canopy N in chapters 3 and 4. One of the research future perspectives of this work would be to attempt and test the estimation of canopy N at global scale using VIs. In order to enable this, data availability, for both canopy N and remote sensing, is essential.

With regards to canopy N data, it is essential that good quality data from all over the globe is available to identify and tune the models. While current datasets on canopy N include data from forests plots well spread over western Europe and North America, regions in the southern hemisphere, i.e. central Africa and south America, suffer from poor data representativity. Ecosystems and vegetation types specific to these regions, e.g. tropical forests, are not well represented in the datasets. It is thus necessary to ensure that data used to calibrate models used to estimate canopy N data include data representing a diversity of ecosystems and conditions to ensure the results are not biased towards some specific ecosystems.

Remote sensing data, on the other hand, are available globally. Recent advances in sensor technical specification enable coverage of the entire earth regularly at high spatial resolution. For example, the MSI sensor aboard the Sentinel 2 satellite from the European Space Agency Copernicus satellite constellation was launched in 2015 and offers coverage over main land surfaces every 5 days and for 14 spectral bands (Drusch et al. 2012). Among those, 3 reflectance bands are designed for vegetation detection as they are placed in the red-edge region of the spectrum, which is important for canopy N estimation with VIs. Next, several spaceborne spectrometers are becoming available in the near future providing data with high spatio-temporal and spectral resolutions. Compared to the spatial resolution of MODIS and MERIS sensors (250 m and 1 km, respectively), which products were used in chapters 3 and

4, Sentinel 2 offers a much higher spatial resolution of 10-60 m. The higher spatial resolution, combined with the red-edge spectral bands, provide the opportunity to estimate canopy N at finer scale and possibly with higher accuracy.

In addition, since remote sensing data are becoming available over regular time steps, they offer the potential to monitor change in canopy over the growing season. GVMs also represent foliage N dynamically over time and show interseasonal variation (Zaehle and Friend 2010). A future perspective would thus be to estimate seasonal canopy N and compare or integrate this temporal variation to the GVM predictions. However, to achieve this at a scale relevant for broad scale vegetation modelling, canopy N data from forest plots would need to be available several times a year and at multiple locations. This would represent a considerable improvement but also would require higher budgets than regular current practice. For example, in the ICP-Forests dataset, the forest plots are sampled once every two years only (Rautio et al. 2016).

Another future development in canopy N estimation would be to include different vegetation types. While chapter 2 tested the possibility to estimate canopy N in grasses, most of the studies presented in this thesis focused on forests. Given that grasslands cover between 30 and 40 % of the terrestrial biosphere and are important C sinks (Gourlez de la Motte et al. 2016; Hossain and Li 2021; Sha et al. 2020), including grasses in a future work estimating canopy N at large scale, provided data are available, would be an interesting development.

Finally, the study presented in chapter 2 laid the ground for the remote sensing based estimation of canopy P and canopy N:P. The results showed that estimation of canopy P and canopy N:P is possible. However, while more research is needed to investigate further if estimating canopy P and canopy N:P with remote sensing is feasible at larger scale and in different ecosystems, this was not the main focus of this thesis. The observation made about N data availability, or lack thereof, is even more valid when it comes to P. Since P, along with N, is an essential nutrient that limits primary productivity (Elser et al. 2007; Vitousek et al. 2010), global estimates of canopy P would be welcome.

6.6 References

- Butler, E.E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K.R., Fazayeli, F., Banerjee, A., Atkin, O.K., Kattge, J., Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K.A., Byun, C., Campetella, G., Cerabolini, B.E.L., Cornelissen, J.H.C., Craine, J.M., Craven, D., de Vries, F.T., Díaz, S., Domingues, T.F., Forey, E., González-Melo, A., Gross, N., Han, W., Hattingh, W.N., Hickler, T., Jansen, S., Kramer, K., Kraft, N.J.B., Kurokawa, H., Laughlin, D.C., Meir, P., Minden, V., Niinemets, Ü., Onoda, Y., Peñuelas, J., Read, Q., Sack, L., Schamp, B., Soudzilovskaia, N.A., Spasojevic, M.J., Sosinski, E., Thornton, P.E., Valladares, F., van Bodegom, P.M., Williams, M., Wirth, C., & Reich, P.B. (2017). Mapping local and global variability in plant trait distributions. *Proceedings of the National Academy of Sciences*, 114, E10937. 10.1073/pnas.1708984114
- Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quere, C., Myneni, R.B., Piao, S., & Thornton, P. (2013). Carbon and other biogeochemical cycles. In T.F. Stocker, Q. D., G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P.M. Midgley (Eds.), *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 465-570). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press
- Davies-Barnard, T., Meyerholt, J., Zaehle, S., Friedlingstein, P., Brovkin, V., Fan, Y., Fisher, R.A., Jones, C.D., Lee, H., Peano, D., Smith, B., Wärlind, D., & Wiltshire, A.J. (2020). Nitrogen cycling in CMIP6 land surface models: progress and limitations. *Biogeosciences*, 17, 5129-5148. 10.5194/bg-17-5129-2020
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meyret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25-36. doi:10.1016/j.rse.2011.11.026
- Elser, J.J., Bracken, M.E.S., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., Ngai, J.T., Seabloom, E.W., Shurin, J.B., & Smith, J.E. (2007). Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters*, 10, 1135-1142. <https://doi.org/10.1111/j.1461-0248.2007.01113.x>
- Exbrayat, J.F., Bloom, A.A., Carvalhais, N., Fischer, R., Huth, A., MacBean, N., & Williams, M. (2019). Understanding the Land Carbon Cycle with Space Data: Current Status and Prospects. *Surveys in Geophysics*, 40, 735-755. 10.1007/s10712-019-09506-2
- Fernández-Martínez, M., Vicca, S., Janssens, I.A., Sardans, J., Luyssaert, S., Campioli, M., Chapin Iii, F.S., Ciais, P., Malhi, Y., Obersteiner, M., Papale, D., Piao, S.L., Reichstein, M., Rodà, F., & Peñuelas, J. (2014). Nutrient availability as the key regulator of global forest carbon balance. *Nature Climate Change*, 4, 471-476. 10.1038/nclimate2177
- Fick, S.E., & Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302-4315. doi:10.1002/joc.5086
- Fleischer, K., Wärlind, D., van der Molen, M.K., Rebel, K.T., Arneth, A., Erisman, J.W., Wassen, M.J., Smith, B., Gough, C.M., Margolis, H.A., Cescatti, A., Montagnani, L., Arain, A., & Dolman, A.J. (2015). Low historical nitrogen deposition effect on carbon sequestration in the boreal zone. *Journal of Geophysical Research: Biogeosciences*, 120, 2542-2561. 10.1002/2015JG002988

- Giuliani, G., Egger, E., Italiano, J., Poussin, C., Richard, J.-P., & Chatenoux, B. (2020). Essential Variables for Environmental Monitoring: What Are the Possible Contributions of Earth Observation Data Cubes? *Data*, 5, 100. <https://doi.org/10.3390/data5040100>
- Gourlez de la Motte, L., Jérôme, E., Mamadou, O., Beckers, Y., Bodson, B., Heinesch, B., & Aubinet, M. (2016). Carbon balance of an intensively grazed permanent grassland in southern Belgium. *Agricultural and Forest Meteorology*, 228-229, 370-383. <https://doi.org/10.1016/j.agrformet.2016.06.009>
- Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangquan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, 12, e0169748. [10.1371/journal.pone.0169748](https://doi.org/10.1371/journal.pone.0169748)
- Hossain, M.L., & Li, J. (2021). Disentangling the effects of climatic variability and climate extremes on the belowground biomass of C3- and C4-dominated grasslands across five ecoregions. *Science of The Total Environment*, 760, 143894. <https://doi.org/10.1016/j.scitotenv.2020.143894>
- Knyazikhin, Y., Lewis, P., Disney, M.I., Möttus, M., Rautiainen, M., Stenberg, P., Kaufmann, R.K., Marshak, A., Schull, M.A., Latorre Carmona, P., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., Yang, Y., & Myneni, R.B. (2013a). Reply to Ollinger et al.: Remote sensing of leaf nitrogen and emergent ecosystem properties. *Proceedings of the National Academy of Sciences*, 110, E2438. [10.1073/pnas.1305930110](https://doi.org/10.1073/pnas.1305930110)
- Knyazikhin, Y., Lewis, P., Disney, M.I., Stenberg, P., Möttus, M., Rautiainen, M., Kaufmann, R.K., Marshak, A., Schull, M.A., Carmona, P.L., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., Yang, Y., & Myneni, R.B. (2013b). Reply to Townsend et al.: Decoupling contributions from canopy structure and leaf optics is critical for remote sensing leaf biochemistry. *Proceedings of the National Academy of Sciences of the United States of America*, 110. [10.1073/pnas.1301247110](https://doi.org/10.1073/pnas.1301247110)
- Knyazikhin, Y., Schull, M.A., Stenberg, P., Möttus, M., Rautiainen, M., Yang, Y., Marshak, A., Latorre Carmona, P., Kaufmann, R.K., Lewis, P., Disney, M.I., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., & Myneni, R.B. (2013c). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences*, 110, E185-E192. [10.1073/pnas.1210196109](https://doi.org/10.1073/pnas.1210196109)
- LeBauer, D.S., & Treseder, K.K. (2008). Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, 89, 371-379. [10.1890/06-2057.1](https://doi.org/10.1890/06-2057.1)
- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., & Martin, M.E. (2016). Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sensing of Environment*, 173, 174-186. [doi:10.1016/j.rse.2015.11.028](https://doi.org/10.1016/j.rse.2015.11.028)
- McNeil, B.E., Read, J.M., & Driscoll, C.T. (2012). Foliar Nitrogen Responses to the Environmental Gradient Matrix of the Adirondack Park, New York. *Annals of the Association of American Geographers*, 102, 1-16. [10.1080/00045608.2011.595654](https://doi.org/10.1080/00045608.2011.595654)
- Moreno-Martínez, Á., Camps-Valls, G., Kattge, J., Robinson, N., Reichstein, M., van Bodegom, P., Kramer, K., Cornelissen, J.H.C., Reich, P., Bahn, M., Niinemets, Ü., Peñuelas, J., Craine, J.M., Cerabolini, B.E.L., Minden, V., Laughlin, D.C., Sack, L., Allred, B., Baraloto, C., Byun, C., Soudzilovskaia, N.A., & Running, S.W. (2018). A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sensing of Environment*, 218, 69-88. <https://doi.org/10.1016/j.rse.2018.09.006>
- Ollinger, S.V., Richardson, A.D., Martin, M.E., Hollinger, D.Y., Frolking, S.E., Reich, P.B., Plourde, L.C., Katul, G.G., Munger, J.W., Oren, R., Smith, M.L., Paw U, K.T., Bolsta, P.V., Cook, B.D., Day, M.C., Martin, T.A., Monson, R.K., & Schmid, H.P. (2008). Canopy nitrogen, carbon assimilation, and albe-

- do in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 19336-19341. doi:10.1073/pnas.0810021105.
- Pacheco-Labrador, J., González-Cascón, R., Pilar Martín, M., & Riaño, D. (2014). Understanding the optical responses of leaf nitrogen in mediterranean holm oak (*Quercus ilex*) using field spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, 26, 105-118. doi:10.1016/j.jag.2013.05.013
- Ramoelo, A., Cho, M., Mathieu, R., Skidmore, A.K., Schlerf, M., Heitkonig, I.M.A., & Prins, H.H.T. (2011). Integrating environmental and in situ hyperspectral remote sensing variables for grass nitrogen estimation in savannah ecosystems. In *34th International Symposium on Remote Sensing of Environment – The GEOSS Era: Towards Operational Environmental Monitoring*
- Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., & Heitkönig, I.M.A. (2012). Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor. *International Journal of Applied Earth Observation and Geoinformation*, 19, 151-162. doi:10.1016/j.jag.2012.05.009
- Ramoelo, A., Skidmore, A.K., Schlerf, M., Heitkönig, I.M.A., Mathieu, R., & Cho, M.A. (2013). Savanna grass nitrogen to phosphorous ratio estimation using field spectroscopy and the potential for estimation with imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, 23, 334-343. doi:10.1016/j.jag.2012.10.008
- Rautio, P., Fürst, A., Stefan, K., Raitio, H., & Bartels, U. (2016). Sampling and Analysis of Needles and Leaves. In UNECE ICP Forests Programme Co-ordinating Centre (Ed.), *Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests* (p. 19 + Annex). Thünen Institute of Forest Ecosystems, Eberswalde, Germany. <http://www.icp-forests.org/manual.htm>. ISBN: 978-3-86576-162-0
- 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 of the United States of America*, 101, 11001. doi:10.1073/pnas.0403588101
- Sardans, J., Alonso, R., Carnicer, J., Fernández-Martínez, M., Vivanco, M.G., & Peñuelas, J. (2016a). Factors influencing the foliar elemental composition and stoichiometry in forest trees in Spain. *Perspectives in Plant Ecology, Evolution and Systematics*, 18, 52-69. doi:10.1016/j.ppees.2016.01.001
- Sardans, J., Alonso, R., Janssens, I.A., Carnicer, J., Veresoglou, S., Rillig, M.C., Fernández-Martínez, M., Sanders, T.G.M., & Peñuelas, J. (2016b). Foliar and soil concentrations and stoichiometry of nitrogen and phosphorous across European *Pinus sylvestris* forests: Relationships with climate, N deposition and tree growth. *Functional Ecology*, 30, 676-689. doi:10.1111/1365-2435.12541
- Sardans, J., Janssens, I.A., Alonso, R., Veresoglou, S.D., Rillig, M.C., Sanders, T.G.M., Carnicer, J., Filella, I., Farré-Armengol, G., & Peñuelas, J. (2015). Foliar elemental composition of European forest tree species associated with evolutionary traits and present environmental and competitive conditions. *Global Ecology and Biogeography*, 24, 240-255. doi:10.1111/geb.12253
- Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L., & Quegan, S. (2017). Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems. *Biogeosciences*, 14, 3401-3429. doi:10.5194/bg-14-3401-2017
- Sha, Z., Bai, Y., Lan, H., Liu, X., Li, R., & Xie, Y. (2020). Can more carbon be captured by grasslands? A case study of Inner Mongolia, China. *Science of The Total Environment*, 723, 138085. <https://doi.org/10.1016/j.scitotenv.2020.138085>

- Tian, Y.C., Yao, X., Yang, J., Cao, W.X., Hannaway, D.B., & Zhu, Y. (2011). Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance. *Field Crops Research*, *120*, 299-310. doi:10.1016/j.fcr.2010.11.002
- Townsend, P.A., Serbin, S.P., Kruger, E.L., & Gamon, J.A. (2013). Disentangling the contribution of biological and physical properties of leaves and canopies in imaging spectroscopy data. *Proceedings of the National Academy of Sciences of the United States of America*, *110*. 10.1073/pnas.1300952110
- Vitousek, P.M., & Howarth, R.W. (1991). Nitrogen limitation on land and in the sea: How can it occur? *Biogeochemistry*, *13*, 87-115. 10.1007/BF00002772
- Vitousek, P.M., Porder, S., Houlton, B.Z., & Chadwick, O.A. (2010). Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen-phosphorus interactions. *Ecological Applications*, *20*, 5-15. <https://doi.org/10.1890/08-0127.1>
- Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A.K., Jones, S., Suarez, L., Woodgate, W., Heiden, U., Heurich, M., & Hearne, J. (2016). Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sensing*, *8*. doi:10.3390/rs8060491
- Wieder, W.R., Cleveland, C.C., Smith, W.K., & Todd-Brown, K. (2015). Future productivity and carbon storage limited by terrestrial nutrient availability. *Nature Geoscience*, *8*, 441. 10.1038/ngeo2413
- Zaehle, S., & Dalmonech, D. (2011). Carbon-nitrogen interactions on land at global scales: Current understanding in modelling climate biosphere feedbacks. *Current Opinion in Environmental Sustainability*, *3*, 311-320. 10.1016/j.cosust.2011.08.008
- Zaehle, S., & Friend, A.D. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, *24*. 10.1029/2009GB003521

Peer-reviewed publications

- Loozen, Y., Rebel, K.T., Karssenber, D., Wassen, M.J., Sardans, J., Peñuelas, J., & de Jong, S.M. (2018). Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index. *Biogeosciences*, 15, 2723-2742. <https://doi.org/10.5194/bg-15-2723-2018>
- Loozen, Y., Karssenber, D., de Jong, S.M., Wang, S., van Dijk, J., Wassen, M.J., & Rebel, K.T. (2019). Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses. *International Journal of Applied Earth Observation and Geoinformation*, 75, 1-14. <https://doi.org/10.1016/j.jag.2018.08.012>
- Loozen, Y., Rebel, K.T., de Jong, S.M., Lu, M., Ollinger, S.V., Wassen, M.J., & Karssenber, D. (2020). Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method. *Remote Sensing of Environment*, 247, 111933. <https://doi.org/10.1016/j.rse.2020.111933>

Curriculum vitae

Yasmina Loozen was born in Liège in 1989. After finishing secondary school in 2008 at the Athénée Royal de Soumagne, she spent a year studying languages. She took Dutch courses in Vught, the Netherlands, and studied English in Bournemouth, the United Kingdom. She then pursued the bachelor degree Bioscience Engineering at Gembloux Agro-Bio Tech, University of Liège from 2009 to 2012. From 2012 to 2014, she pursued the master degree Bioscience Engineering in Environmental Sciences and Technologies at Gembloux Agro-Bio Tech, University of Liège. During this master degree, she had the opportunity to study one semester at the Department of Physical Geography and Ecosystem Sciences of Lund University, Sweden, with the Erasmus program. During this exchange, she focused on the Carbon cycle and Ecosystem Dynamics. Her master thesis project focused on the design of a biofilter and testing of biofilter media for an aquaponic application at the Unité de Phytopathologie Intégrée et Urbaine of Gembloux Agro-Bio Tech, University of Liège. After graduating her master with Magna Cum Laude, she started her PhD in January 2015 at the Geosciences Faculty of Utrecht University as a joined project between the Department of Physical Geography and the Copernicus Institute of Sustainable Development. She worked under the supervision of Karin Rebel, Derek Karssenbergh, Martin Wassen and Steven de Jong. During her PhD research, she combined statistical data analysis, laboratory experiment and machine learning. She collaborated with scientific researchers at the University of New Hampshire, USA, CREAM, Spain and Max Planck Institute for Biogeochemistry, Germany. She is currently working as a researcher at ISSeP, the Walloon Scientific Institute of Public Service, within the Remote Sensing and Geodata Unit.

