Optical remote sensing of vegetation canopy properties from local to global scales

Eleanor Thomson

Supervised by:

Yadvinder Malhi (University of Oxford),

and

Christopher E. Doughty (Northern Arizona University)

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Oxford

Michaelmas 2021











Abstract

We are in the midst of a new era of Earth Observation for vegetation monitoring. Recent enhancements in the spectral, spatial and temporal resolution, as well as availability, of optical sensors provides opportunities to significantly advance the retrieval of vegetation properties at multiple scales. However, directed research linking optical imagery to field-collected vegetation measurements is required to optimally harness current and future advances in remote sensor technologies.

This thesis aims to investigate the optical remote sensing of plant functional traits at four scales: the leaf, unmanned aerial vehicle (UAV), aircraft and satellite level. This thesis has four objectives, each of which are addressed through a research article.

The main objective of the first article, entitled '**Multiscale mapping of plant functional groups and plant traits in the High Arctic using field spectroscopy, UAV imagery and Sentinel-2 data**', is to evaluate the use of multispectral UAV imagery for mapping biochemical leaf traits in high Arctic tundra. The High Arctic is an ideal environment for establishing a baseline potential for mapping leaf traits using UAV imagery, as species richness is low and vegetation cover lacks the structure and confounding effects of a forest canopy. The article starts by using field spectroscopy to classify plant species and quantify biochemical plant traits (water content, nitrogen (N), phosphorus (P) and C:N ratios). At the UAV level, it classifies three plant functional types (mosses, graminoids and dwarf shrubs) at 72% accuracy and generates maps of plant chemistry. (r^2 = 0.44–0.54, RMSEmean = 26%–48%). Lastly, the UAV results were used to explore two methods of upscaling plant water content to the wider landscape using Sentinel-2 imagery.

The objective of the second article, entitled '**Mapping the Leaf Economic Spectrum across West African Tropical Forests Using UAV-Acquired Hyperspectral Imagery**' is to evaluate the use of hyperspectral UAV imagery for mapping biochemical leaf traits across a species-rich, structurally complex forest canopy in Ghana, West Africa. Similarly, the article starts by using field spectroscopy to estimate six biochemical plant traits (leaf mass per area (LMA), N, P, calcium (Ca) and magnesium (Mg)). At the airborne level, four of the six leaf traits demonstrated weak (p < 0.10) correlations with the UAV-collected

hyperspectral imagery: r^2 = 0.25 (LMA), r^2 = 0.22 (N), r^2 = 0.22 (P), and r^2 = 0.25 (Ca). Mapping the leaf economic spectrum (LMA, N and P) across a precipitation gradient reflected carbon cycling data, and revealed precipitation and substrate as co-dominant drivers of trait distributions and relationships.

The objective of the third article, entitled '**Mapping photosynthetic capacity in a Bornean tropical forest from airborne and satellite imagery: challenges and recommendations**', expands focus from mapping biochemical leaf traits to mapping photosynthetic capacity ($Vcmax_{25}$ and $Jmax_{25}$) across a forest disturbance gradient in Malaysian Borneo using aircraft-collected hyperspectral imagery and Sentinel-2 imagery. In addition, it explores increased image quality issues associated with expanded spatial coverage. Aircraft-collected hyperspectral imagery was found to exhibit poor radiometric calibration and Sentinel-2 imagery was found to suffer from processing errors related to cloud masking and inconsistent surface reflectance values, especially over humid tropical forests. The strongest model performance (r^2=0.36 and 0.43 for $Vcmax_{25}$ and $Jmax_{25}$ respectively) was produced by tuning the high spectral and spatial resolution of the aircraft-collected hyperspectral imagery with the improved signal-to-noise ratio of a Sentinel time-series composite, demonstrating the importance of diagnosing and minimising data quality issues.

The last article entitled '**Global mapping of the photosynthetic capacity of tropical forests**', combines Sentinel-2A imagery, environmental datasets and photosynthesis data from 35 tropical forest GEM plots to create a pan-tropical map of $Vcmax_{25}$ and $Jmax_{25}$. To the best of my knowledge, this is the first research that uses Sentinel-2 imagery to map tropical forest $Vcmax_{25}$ and $Jmax_{25}$ globally. Models captured 38% and 57% of variation in $Vcmax_{25}$ and $Jmax_{25}$ across lowland forest, and 62% and 61% of variation across lowland-and-montane forest. In lowland-only forest, mean climatic water deficit (MCWD) and solar radiation were most important predictors of $Vcmax_{25}$ and $Jmax_{25}$ respectively. In lowland-and-montane forest, the strongest predictor of both $Vcmax_{25}$ and $Jmax_{25}$ was mean annual temperature (MAT). At intercontinental scales, $Vcmax_{25}$ and $Jmax_{25}$ values were mainly controlled by climate. At the regional level, trait distributions were shaped by soil type. At local levels, trait patterns reflected spectral variation arising from topography and canopy structure.

Overall, this thesis offers new methodologies to measure and monitor tropical forest functional diversity and productivity. It builds from a simple ecosystem to a floristically and structurally complex tropical forest, and explores the trade-offs of expanding spatial coverage and decreasing spectral resolution for the retrieval of vegetation canopy traits. Alongside, it highlights current limitations of optical remote sensing, and presents best-practice recommendations for future imaging studies. The results are pertinent for high resolution, low cost mapping of tundra and tropical forest ecosystems (and potentially for other terrestrial ecosystems), and for improving the representation of vegetation properties, especially photosynthetic capacity, in the next generation of Earth System Models.

Acknowledgements

This research would not have been possible without the support of several people. First and foremost, I would like to thank my supervisors Yadvinder Malhi and Chris Doughty. To Yadvinder, thank you for answering the very first email I sent you in 2012, and for all the guidance, support and opportunities you have so generously offered me throughout the course of my academic career. I always left your office feeling more inspired and optimistic about my research, and this work has benefited greatly from your experience and vision.

To Chris, I am immensely grateful for introducing me to the field of remote sensing, and imaging spectroscopy in particular. Thank you for sharing your knowledge with me and for the many collaborations you have offered. In particular, thank you for welcoming me to your inaugural lab at Northern Arizona University and for all the time and support you gave me during that year, and since. The scope of this thesis reflects many of your original ideas and creativity.

At Northern Arizona University, I am also extremely grateful to Andrew Abraham, Chase Ridenour and Tomos Prys-Jones for their hospitality and friendship. My visits would have been much more difficult without their help, and much less enjoyable without their company.

Throughout this thesis, I have been fortunate to receive a number of opportunities that have strengthened this research. Thank you to the Plant Functional Trait Course 4 (PFTC) team, especially Vigdis Vandvik and Brian Enquist for organising the logistics of fieldwork in Svalbard that formed the primary data collection of Chapter 5. Very special thanks to all the members of the remote sensing group who helped collect the data, especially Marc Macias-Fauria, whose input and guidance throughout this project was invaluable, and Marcus Spiegel who stuck with this project throughout the many ups and downs, and taught me the foundations of spatial analysis in R.

Thank you to the Wageningen University and Research Unmanned Aerial who collected the UAV hyperspectral imagery in Chapter 6, and especially to Juha Suomalainen who

taught me the basics of working with it. Thank you also to David Coomes' and his research group in Cambridge, particularly Tom Swinfield and Boris Bongalov, who shared the Borneo aircraft imagery with me and answered my many questions so patiently.

Lastly, this thesis would not have been possible without primary financial support from a NERC DTP award (NE/L002612/1), and several Oriel College Travel Grants that allowed me to undertake fieldwork and present my research at conferences.

And finally, to my family and dear friends who continue to support me, even though I have neglected them since the start of this PhD, I am grateful for their love, patience and understanding.

Table of Contents

Abstract **2**
Acknowledgements **5**
List of abbreviations **9**
Thesis overview **10**

Chapter 1: Introduction **11**
1.1. Thesis rational **12**
1.2. Thesis aims and objectives **19**
1.3. References **21**

Chapter 2: History and context **27**
2.1. Airborne mapping of biochemical traits **28**
2.2. Optical remote sensing of process-based traits **33**
2.3. Satellite-based mapping of photosynthetic capacity **36**
2.4. References **38**

Chapter 5: Multiscale mapping of plant functional groups and plant traits in the High Arctic using field spectroscopy, UAV imagery and Sentinel-2 data **44**
5.1. Chapter statement **45**
5.2. Author contributions **46**
5.3. Abstract **47**
5.4. Introduction **48**
5.5. Methods **52**
5.6. Results **64**
5.7. Discussion **76**
5.8. Conclusion **82**
5.9. Acknowledgments **83**
5.10. Data availability statement **83**
5.11. Funding **83**
5.12. References **84**
5.13. Supplementary material **94**

Chapter 6: Mapping the Leaf Economic Spectrum across West African Tropical Forests Using UAV-Acquired Hyperspectral Imagery. **99**
6.1. Chapter statement **100**
6.2. Author contributions **101**
6.3. Abstract **102**
6.4. Introduction **103**
6.5. Materials and Methods **107**
6.6. Results **116**
6.7. Discussion **127**
6.8. Conclusion **132**
6.9. Acknowledgments **133**
6.10. Funding **133**
6.11. References **134**
6.12. Supplementary material **141**

Chapter 7: Mapping photosynthetic capacity in a Bornean tropical forest from airborne and satellite imagery: challenges and recommendations **142**
7.1. Chapter statement **143**
7.2. Author contributions **144**
7.3. Abstract **145**
7.4. Introduction **146**
7.5. Materials and methods **150**
7.6. Results **162**
7.7. Discussion **175**
7.8. Conclusion **180**
7.9. References **181**
7.10. Supplementary material **187**

Chapter 8: Global mapping of the photosynthetic capacity of tropical forests **195**
8.1. Chapter statement **196**
8.2. Author contributions **197**
8.3. Abstract **198**
8.4. Introduction **199**
8.5. Materials and methods **204**
8.6. Results **217**
8.7. Discussion **229**
8.8. Conclusion **236**
8.9. References **238**
8.10. Supplementary material **247**

Chapter 9: Conclusion **265**
9.1. Thesis synthesis **266**
9.2. Main implication and future research needs **269**
9.3. References **274**

List of abbreviations

AOT: Aerosol optical thickness
AVIRIS: Airborne Visible / Infrared Imaging Spectrometer
BOA: Bottom-of-atmosphere
BRDF: Bidirectional reflectance distribution function
Ca: Leaf calcium concentration
CAO: Carnegie Airborne Observatory
CNN: Convolutional neural network
DBH: Diameter at breast height
DGVM: Dynamic global vegetation model
DLS: Downwelling light sensor
DSM: Digital surface model
ESA: European Space Agency
FWHM: Full width at half maximum
GCP: Ground control point
GEM: Global Ecosystems Monitoring
GNSS: Global navigation satellite system
GPP: Gross primary productivity
ITCZ: International tropical convergence zone
ITEX: International Tundra Warming Experiment
K: Leaf potassium concentration
LAD: Leaf angle distribution
LAI: Leaf area index
LES: Leaf economic spectrum
LiDAR: Light detection and ranging
LMA: Leaf mass per area
MAT: Mean annual temperature
MCWD: Mean climatic water deficit
Mg: Leaf magnesium concentration
N: Leaf nitrogen concentration
NASA: National Aeronautics and Space Administration
NDVI: Normalized difference vegetation index
NIR: Near-infrared
NPP: Net primary productivity
P: Leaf phosphorus concentration
PFG/T: Plant functional group/type
PLSR: Partial least squares regression
RGB: Red, green, blue
RMSE: Root mean square error
SCL: Scene classification layer
SfM: Structure-from-motion
SWIR: Shortwave infrared
TBM: Terrestrial biosphere model
TOA: Top-of-atmosphere
UAV: Unmanned aerial vehicle
VNIR: Visible to near-infrared
VSWIR: Visible – shortwave infrared
WV: Water vapour

Thesis Overview

This thesis is article-based and presented according to the requirements for the University of Oxford's DPhil degree in Geography and the Environment. The thesis is organised into nine chapters. **Chapter 1** introduces the thesis and sets out the specific aims and research questions it will address. **Chapter 2** contextualises the thesis in the relevant literature and identifies the key research areas it builds upon. **Chapter 3** describes the main field sites where the research took place. **Chapter 4** describes the main spectral collection and machine learning methods used in the thesis. **Chapters 5, 6, 7** and **8** address the objectives of the thesis in four research articles. Finally, **Chapter 8** discusses the main conclusions and implications of the research.

Chapter 1:

Introduction



Photo: Eleanor Thomson

1.1 Thesis rationale

1.1.1. An introduction to plant functional traits

The terrestrial biosphere plays a central role in the functioning of the global climate system. It accounts for 60% of Earth's primary productivity (Beer *et al*. 2010), has sequestered one third of anthropogenic CO_2 emissions since 1960 (Le Quere *et al*. 2014), and strongly regulates regional and global climate through carbon and water cycling, and albedo effects (Malhi *et al*. 2002; Foley *et al*. 2003; Meir *et al*. 2006; Beer *et al*. 2010; Pearson *et al*. 2013). These globally important roles are mediated through a simple structure: the leaf. Leaves have a range of properties, or traits. Leaf functional traits refer to structural, chemical and biological characteristics that determine factors such as a plant's growth rate, lifespan, ecological strategy, metabolic activity and tolerance of pests and hazards (Kattge *et al*. 2011). Together the range of trait values in an ecosystem links species richness to ecosystem function and determines how plant communities relate and respond to their environment (McGill *et al*. 2006; Westoby and Wright, 2006). As leaf traits both reflect and predict ecosystem function, there is currently much interest in using plant traits to better understand the drivers of ecological processes, and improve predictions of ecosystem response to environmental change (Lavorel and Garnier *et al*. 2002; Sunding and Goldstein, 2008; Funk *et al*. 2016).

Strong mechanistic relationships between leaf traits form the basis of trait-based ecology (Díaz *et al*. 2014; Díaz *et al*. 2016). In particular, leaf traits demonstrate a direct link to ecosystem function via the Leaf Economic Spectrum (LES). The LES describes an axis of six main leaf traits (leaf mass per area (LMA), leaf nitrogen (N), leaf phosphorus (P), leaf longevity, photosynthetic assimilation rate, and dark respiration rate) that covary in predictable ways and underpin plant resource acquisition–defence trade-offs (Reich *et al.* 1997; Wright *et al*. 2004). On one end of the spectrum, species opt for a quick return on investments by producing 'cheap', short-lived leaves with high N and P concentrations that facilitate high photosynthesis rates and fast growth. These species are commonly found in high-fertility environments where short-lived leaves (because of abscission and herbivory) maintain relatively fast rates of nutrient cycling (Santiago, 2007; Cornwell *et al*. 2008). On the other end of the spectrum, species located in lower precipitation and

nutrient regimes tend to trade rapid growth for persistence and invest in tough, well-protected leaves, with higher LMA values and lower N and P concentrations. They produce leaves that can better withstand herbivory and physical hazards, leading to increased leaf longevity and lower rates of photosynthesis and nutrient cycling (Poorter and Markesteijn, 2008; Hoffmann *et al*. 2015). The LES has been shown to be a good indicator of plant performance (Poorter and Bongers, 2006; Poorter *et al*. 2014), with a species' position on the spectrum describing a plant's overall strategy (quick growing but ephemeral *versus* slow growing but enduring). While some deviations have been reported at a local scale (Messier *et al*. 2010; Wright and Sutton-Grier, 2012; Feng and Dietze, 2013), this spectrum has been shown to operate similarly across a wide variety of species and biomes, with important implications for global dynamic vegetation and climate models (Sakschewski *et al*. 2015).

1.1.2 The plant trait data deficit

Despite the fact that functional traits have been found to form the basis of a more quantitative and predictive ecology, a major limitation to this research has been a lack of spatially extensive trait data. While global databases such as GLOPNET (Wright *et al*. 2004, 2005) and TRY (Kattge *et al*. 2011, 2020) have made great progress in addressing the trait data deficit over the last few decades, large gaps and sampling biases in these databases still continue to pose a challenge (Cornwell *et al.* 2019). Spatially-explicit trait data are especially important for improving the representation of vegetation in terrestrial biosphere models (TBMs) to generate robust projections of regional and global change (McMahon *et al*. 2011). Most TBMs simulate vegetation though a small number of plant functional types (PFTs), based on classic climate limitations such as temperature, precipitation and length of growing season (e.g. Box, 1996; Bonan *et al*. 2002; Sitch *et al*. 2003). Each PFT is parametrized with the same ecological values, with no variation represented within a PFT, or gradual transitions represented between different PFTs.

As a large number of observations are needed for sufficient PFT classification (Wang and Price, 2007), not only would increased trait data help to delineate PFGs, and assign accurate parameter values, but spatially-continuous data could fuel a revolution in vegetation models, away from simulating vegetation using small number of discrete,

homogenous PFTs, towards a more spatially-nuanced, trait-based representation of vegetation. Experiments using a small number of trait-based dVGMs (dynamic vegetation global models) have reported more realistic representations of functional diversity and climate-vegetation interactions (Scheiter *et al.* 2013; Sakschewski *et al.* 2015). Thus, comprehensive biological data is the most significant area where substantial advances could be made in climate-biosphere modelling (McMahon *et al*. 2011).

Nowhere is this lack of trait data more severe than the tropics (Jetz *et al*. 2016). Tall trees, high species diversity and the labour-intensive nature of collecting trait data, combined with the inaccessibility of many tropical regions, means measuring and monitoring functional traits in tropical areas is a time-consuming and expensive process. Yet tropical forests account for one-third of Earth's terrestrial primary productivity (Beer *et al*. 2010) and store 40-50% of global vegetation carbon (Pan *et al*. 2011; Houghton *et al*. 2012). Thus, determining the functional diversity of tropical forest canopies is critical to predicting future climate–biosphere interactions.

1.1.3 Remote sensing of leaf traits

Remote sensing would be an ideal method for generating spatially-continuous trait data and addressing the tropical forest data deficit. Using high-resolution spectroscopy to measure plant foliar traits has a relatively long scientific history (Gates *et al*. 1965; Knipling, 1970). Spectrometers measure light reflected from a leaf surface in narrow, contiguous spectral bands across a broad wavelength (visible-near infrared (VNIR: 400–1100 nm), or visible-short-wave infrared (VSWIR: 400–2500 nm)). Specific spectral reflection and absorption features have been directly linked to a leaf's concentration of cellulose, lignin, chlorophylls, N, starch, oil, proteins and water, listed by Curran (1989) in a well-known synthesis. For example, the measurement of N, bound in chlorophyll and proteins, is associated with the bandwidths between 460–480, 650–670 and 1510–1600 nm (Kokaly, 2001; Smith *et al*. 2003), while LMA can be indirectly estimated using variations in leaf thickness and leaf water content expressed across the 700–1300 nm range (Asner *et al*. 2011a; Serbin *et al*. 2019).

In the last decade, forefronted by Asner *et al*., significant headway has been made in the remote sensing of tropical forest leaf traits, both at the individual leaf level and at the canopy level (Asner and Martin, 2008; Asner *et al*. 2011b, 2015, 2016, 2017; Doughty *et al*. 2017; Nunes *et al*. 2019). Using partial least squares regression analysis (PLSR), Asner *et al*. have developed relationships between leaf spectra and a suite of biochemical and structural leaf traits (Asner and Martin, 2008; Asner *et al*. 2011b, 2015). The PLSR technique is useful because it does not focus on a small number of wavelengths but uses the entire leaf spectral continuum in its analysis. Using PLSR, relationships have been also found between leaf spectra and macronutrients, such as P, K, Ca and Mg, that have no direct expression in the VNIR or VSWIR, but often demonstrate correlations with other traits (Asner *et al*. 2011b, 2012a; Stein *et al*. 2014; Gökkaya *et al*. 2015). Similarly, relationships between spectra and 'process-based' traits, such as photosynthetic capacity, have been found at the leaf level, generally driven through correlations with N and chlorophyll (Doughty *et al*. 2011; Serbin *et al*. 2012; Ainsworth *et al*. 2014; Barnes *et al*. 2017; Wu *et al*. 2021).

While handheld spectroscopy cannot generate spatially-continuous maps of leaf traits, handheld spectroscopy is often used as a baseline for determining the strength of spectra-trait relationships. Unlike during airborne spectroscopy, leaf surfaces are kept flat to minimize scatter, a powerful artificial light source is used to maximize photon energy at each wavelength, and multiple spectra are integrated during a measurement to maximize signal-to-noise ratio. If strong spectra-trait relationships cannot be determined in these 'laboratory' conditions, there is little point scaling up to the airborne level, as airborne spectroscopy suffers from decreased signal-to-noise ratio arising from variable leaf angle, weaker and inconsistent solar illumination, and singular spectral measurements.

1.1.4 Manned aircraft mapping of leaf traits

At the airborne level, physical and chemical leaf traits have been mapped across large swathes of tropical forest using high resolution imaging spectrometers mounted onboard aircraft. The Carnegie Airborne Observatory (Asner *et al*. 2007; 2012b) has led the charge in this area, revolutionizing landscape-level understanding of leaf traits over Hawaii, Peru, and lately Borneo (Asner *et al.* 2008a, 2008b, 2016; 2017; Martin *et al.* 2018). These

studies have revealed the wide trait variation present across tropical forests - regions which have hitherto been treated as a single PFT in Earth System Models.

Yet while manned airborne mapping studies have highlighted the importance of expanding trait-ecology from spatially-limited plot studies, they suffer from a couple of key limitations (**Table 1.1**). Their flights are typically singular, providing only a single snapshot in time, and lacking information on the seasonal and inter-annual variation in leaf traits (Rossatto *et al.* 2013; Wagner *et al.* 2017; Nunes *et al*. 2019). Few platforms of this type exist and the average researcher does not have practical or financial access to such equipment, limiting spatial coverage and innovation in this area. In addition, the ability to reproduce experiments and observations is at the heart of science, with growing alarm about irreproducible results (*Nature Special Issue*, 2014). Thus, in order to verify these studies through scientific corroboration, the democratization of these methods is essential.

	Mapping method		
	UAV	Aircraft	Sentinel-2
Advantages	• High spatial resolution (~5cm). • Relatively low-cost system. • Available to a wide range of researchers. • Easy and flexible to deploy. • Few issues with cloud cover and haze.	• Medium spatial resolution (~50cm). • High spatial coverage. • Most systems are hyperspectral.	• Global coverage. • High revisit time (5 days). • Freely available imagery.
Disadvantages	• Small spatial coverage. • 'Off-the-shelf' systems are generally multispectral. • Few hyperspectral systems available.	• Expensive to build and run. • Only available to a small number of researchers. • Imagery is rarely made public. • Few repeat flights. • Some issues with cloud cover and haze.	• Low spatial resolution (10-20m). • Multispectral imagery only. • Issues with cloud cover and haze, especially over tropical regions.

Table 1.1: Table comparing the advantages and disadvantages of three mapping methods: UAV, aircraft, and satellite (Sentinel 2).

1.1.5 UAV mapping of leaf traits

Many of the weaknesses of aircraft-based remote sensing are the strengths of a UAV mapping approach (**Table 1.1**). In comparison to aircraft, UAVs are accessible to a wide array of researchers. They can operate at short notice, without a planned schedule, and at low cost. Although their flight duration is much shorter than that of an aircraft, limiting their area of data collection, they can fly slower, at much lower altitudes, and acquire centimetre to decimetre spatial resolution data, which is less affected by atmospheric conditions (Pajares, 2015). Using UAVs to map and monitor tropical forest canopy traits would have distinct advantages over aircraft in that they would enable a wider proportion of the research community to quantify canopy traits, increasing spatial coverage, innovation, and scientific corroboration in this area. Especially in Arctic literature, there is growing appreciation of the limits of plot-based field sampling, which is biased to a few accessible research areas and fails to capture a large proportion of ecological diversity (Metcalfe *et al*. 2018). UAVs could be deployed on multiple occasions to monitor temporal changes, and can be deployed quickly after a disturbance event, such as a drought or fire, to measure its effects. Additionally, complementarities in resolution between UAV and satellite acquisition, can help bridge the 'scale gap' between field-sampled traits and coarse-scale satellite imagery, and enable field traits to be scaled up over wider areas (Assman *et al*. 2020; Revill *et al*. 2020).

1.1.6 Satellite mapping of leaf traits

As an alternative to collecting private manned and UAV imagery, NASA and ESA provide freely-available satellite imagery. High spatial resolution is traded for global coverage and a frequent revisit time (**Table 1.1**). Landsat and MODIS imagery have been used to create local to global maps of biochemical leaf traits, often through the use of vegetation indices that are based on two or more spectral bands and are designed to highlight specific vegetation characteristics (Running *et al*. 2004; Houborg *et al*. 2007; Wu *et al*. 2010). These vegetation indices rely on simple statistical relationships between leaf traits and spectra, but often saturate over tropical forests (Propastin *et al*. 2012; Cleveland *et al*. 2015). More recently, Sentinel-2 imagery, combined with environmental datasets and more flexible machine learning methods, has been used to scale up field-measurements and create pan-tropical maps of biochemical leaf traits (Aguirre-Gutiérrez *et al*. 2021). To

the best of our knowledge however, Sentinel-2 data has never been used to map process-based traits, such as photosynthetic capacity.

This project aims to build on and diversify the work discussed above in three ways:

1. Demonstrating the use of Unmanned Aerial Vehicles (UAVs) as a low-cost, high resolution tool for mapping leaf functional traits.
2. Expanding a focus from biochemical leaf traits to mapping process-based traits, particularly photosynthetic capacity.
3. Using Sentinel-2 imagery to create a pan-tropical map of photosynthetic capacity.

These aims are addressed through four chapters. **Chapter 5** evaluates the use of multispectral UAV imagery for mapping biochemical leaf traits in high Arctic tundra. The High Arctic is an ideal environment for establishing the baseline potential of mapping leaf traits using UAV imagery, as species richness is much lower and vegetation cover lacks the structure and confounding effects of a tropical forest canopy. **Chapter 6** builds on **Chapter 5** and evaluates the use of hyperspectral UAV imagery for mapping biochemical leaf traits across a species-rich, structurally complex forest canopy in Ghana, West Africa. **Chapter 7** expands a focus from mapping biochemical leaf traits to mapping photosynthetic capacity ($Vcmax_{25}$ and $Jmax_{25}$) across a forest disturbance gradient in Malaysian Borneo. It uses aircraft-collected hyperspectral imagery, and Sentinel-2 imagery, to expand spatial coverage to the regional landscape and explores increased image quality issues associated with inconsistent illumination and atmospheric effects. **Chapter 8** builds on the regional analysis of **Chapter 7** to create a pan-tropical map of $Vcmax_{25}$ and $Jmax_{25}$ by combining Sentinel-2 imagery, environmental datasets and photosynthesis data from 35 tropical forest GEM plots.

1.2. Thesis aims and objectives

The overarching aims of this thesis were addressed through four research articles, each of which had a guiding objective.

Objective 1: To evaluate the use of multispectral UAV imagery for mapping biochemical leaf traits when species richness and vegetation structure are minimised.

To fulfil this objective, the following research aims were addressed in Svalbard, High Arctic:

1. To develop predictive models of species identity and biochemical plant traits (water, N, P and C:N concentrations) using field spectroscopy.
2. To develop predictive models of species identity and the same plant traits using multispectral UAV spectra.
3. To apply the UAV models to UAV imagery to derive spatially continuous maps of plant species and their corresponding trait values.
4. To explore methods of upscaling the UAV maps to the wider landscape using Sentinel-2 imagery are most effective.

Objective 2: To evaluate the use of hyperspectral UAV imagery for mapping biochemical leaf traits across a species-rich, structurally complex forest canopy.

To fulfil this objective, the following research questions were addressed across a precipitation gradient in Ghana, West Africa:

1. Can field spectroscopy predict LMA, N, P, K, Ca and Mg in sun and shade leaves at the individual leaf level?
2. Can the same foliar traits be quantified with similar accuracy at the tree crown level using UAV-collected hyperspectral imagery?
3. Does the spatial distribution of traits and their relationships subscribe to general LES theory?

Objective 3: To explore the use of aircraft-collected hyperspectral imagery, and Sentinel-2 imagery, to develop predictive models of photosynthetic capacity.

To fulfil this objective, the following research questions were addressed across a forest disturbance gradient in Sabah, Malaysian Borneo:

1. What are the most likely causes of cross-sensor disagreements between airborne hyperspectral imagery and Sentinel imagery taken in the same season?
2. What are the most likely causes of internal sensor disagreements between three single Sentinel images, and a Sentinel time-series composite, taken in the same season?
    i) Inaccuracy in Sentinel's cloud detection algorithm (SCL)?
    ii) Inaccurate surface reflectance values?
3. Which type of spectral imagery produces the strongest PLSR model for estimating photosynthetic capacity ($Vcmax_{25}$ and $Jmax_{25}$) across a forest disturbance gradience in Malaysian Borneo?
4. Does combining the high spectral and spatial resolution of the airborne hyperspectral imagery with the stronger signal-to-noise ratio of a Sentinel time-series composite improve estimation of $Vcmax_{25}$ and $Jmax_{25}$?

Objective 4: To use Sentinel-2 imagery to create a pan-tropical map of photosynthesis capacity.

To fulfil this objective, the following research questions were addressed using datasets collected across 6 pan-tropical sites (Australia, Brazil, Gabon, Ghana, Malaysia and Peru):

1. How does photosynthetic capacity ($Vcmax_{25}$ and $Jmax_{25}$) vary between 6 tropical forest sites and across environmental gradients at each site?
2. Which spectral and environmental variables are most important in predicting $Vcmax_{25}$ and $Jmax_{25}$ information pan-tropically?
3. How does the magnitude and spatial patterns of $Vcmax_{25}$ and $Jmax_{25}$ vary within, and between, lowland and montane tropical forest ecosystems?

1.4 References

Aguirre-Gutiérrez, J., Rifai, S., Shenkin, A., Oliveras, I., Bentley, L.P., Svátek, M., Girardin, C.A.J., Both, S., Riutta, T. and Berenguer, E. (2021) Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data. *Remote Sensing of Environment,* 252, pp. 112122.

Ainsworth, E.A., Serbin, S.P., Skoneczka, J.A. and Townsend, P.A. (2014) Using leaf optical properties to detect ozone effects on foliar biochemistry. *Photosynthesis Research,* 119 (1), pp. 65-76.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J.W. and Field, C.B. (2007) Carnegie airborne observatory: in-flight fusion of hyperspectral imaging and waveform light detection and ranging for three-dimensional studies of ecosystems. *Journal of Applied Remote Sensing,* 1 (1), pp. 013536.

Asner, G.P., Jones, M.O., Martin, R.E., Knapp, D.E. and Hughes, R.F. (2008a) Remote sensing of native and invasive species in Hawaiian forests. *Remote Sensing of Environment,* 112 (5), pp. 1912-1926.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J. and Hughes, R.F. (2008b) Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sensing of Environment,* 112 (5), pp. 1942-1955.

Asner, G.P., Martin, R.E., Tupayachi, R., Emerson, R., Martinez, P., Sinca, F., Powell, G.V.N., Wright, S.J. and Lugo, A.E. (2011a) Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. *Ecological Applications,* 21 (1), pp. 85-98.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C., Carranza, L., Martinez, P., Houcheime, M., Sinca, F. and Weiss, P. (2011b) Spectroscopy of canopy chemicals in humid tropical forests. *Remote Sensing of Environment,* 115 (12), pp. 3587-3598.

Asner, G.P., Martin, R.E. and Suhaili, A.B. (2012a) Sources of canopy chemical and spectral diversity in lowland Bornean forest. *Ecosystems,* 15 (3), pp. 504-517.

Asner, G.P., Knapp, D.E., Boardman, J., Green, R.O., Kennedy-Bowdoin, T., Eastwood, M., Martin, R.E., Anderson, C. and Field, C.B. (2012b) Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. *Remote Sensing of Environment,* 124, pp. 454-465.

Asner, G.P., Martin, R.E., Anderson, C.B. and Knapp, D.E. (2015) Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sensing of Environment,* 158, pp. 15-27.

Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E. and Vaughn, N. (2016) Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proceedings of the National Academy of Sciences,* 113 (28), pp. E4043-E4051.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn, N.R. and Llactayo, W. (2017) Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science,* 355 (6323), pp. 385-389.

Asner, G.P. and Martin, R.E. (2008) Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment,* 112 (10), pp. 3958-3970.

Assmann, J.J., Myers-Smith, I., Kerby, J.T., Cunliffe, A.M. and Daskalova, G.N. (2020) Drone data reveal heterogeneity in tundra greenness and phenology not captured by satellites. *Environmental Research Letters,* 15 (12), pp. 125002.

Barnes, M.L., Breshears, D.D., Law, D.J., Van Leeuwen, W.,J.D., Monson, R.K., Fojtik, A.C., Barron-Gafford, G. and Moore, D.J.P. (2017) Beyond greenness: detecting temporal changes in photosynthetic capacity with hyperspectral reflectance data. *PLoS One,* 12 (12), pp. e0189539.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M.A., Baldocchi, D. and Bonan, G.B. (2010) Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science,* 329 (5993), pp. 834-838.

Bonan, G.B., Levis, S., Kergoat, L. and Oleson, K.W. (2002) Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochemical Cycles,* 16 (2), pp. 5-1.

Box, E.O. (1996) Plant functional types and climate at the global scale. *Journal of Vegetation Science,* 7 (3), pp. 309-320.

Cleveland, C.C., Taylor, P., Chadwick, K.D., Dahlin, K., Doughty, C.E., Malhi, Y., Smith, W.K., Sullivan, B.W., Wieder, W.R. and Townsend, A.R. (2015) A comparison of plot-based satellite and Earth system model estimates of tropical forest net primary production. *Global Biogeochemical Cycles,* 29 (5), pp. 626-644.

Cornwell, W.K., Cornelissen, J.H.C., Amatangelo, K., Dorrepaal, E., Eviner, V.T., Godoy, O., Hobbie, S.E., Hoorens, B., Kurokawa, H. and Pérez-Harguindeguy, N. (2008) Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecology Letters,* 11 (10), pp. 1065-1071.

Cornwell, W.K., Pearse, W.D., Dalrymple, R.L. and Zanne, A.E. (2019) What we (don't) know about global plant diversity. *Ecography,* 42 (11), pp. 1819-1831.

Curran, P.J. (1989) Remote sensing of foliar chemistry. *Remote Sensing of Environment,* 30 (3), pp. 271-278.

Diaz, S., Hodgson, J.G., Thompson, K., Cabido, M., Cornelissen, J.H.C., Jalili, A., Montserrat-Martí, G., Grime, J.P., Zarrinkamar, F. and Asri, Y. (2004) The plant traits that drive ecosystems: evidence from three continents. *Journal of vegetation science,* 15 (3), pp. 295-304.

Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C. and Prentice, I.C. (2016) The global spectrum of plant form and function. *Nature,* 529 (7585), pp. 167-171.

Doughty, C.E., Asner, G.P. and Martin, R.E. (2011) Predicting tropical plant physiology from leaf and canopy spectroscopy. *Oecologia,* 165 (2), pp. 289-299.

Doughty, C.E., Santos-Andrade, P., Goldsmith, G.R., Blonder, B., Shenkin, A., Bentley, L.P., Chavana-Bryant, C., Huaraca-Huasco, W., Díaz, S. and Salinas, N. (2017) Can leaf spectroscopy predict leaf and forest traits along a Peruvian tropical forest elevation gradient? *Journal of Geophysical Research: Biogeosciences,* 122 (11), pp. 2952-2965.

Feng, X. and Dietze, M. (2013) Scale dependence in the effects of leaf ecophysiological traits on photosynthesis: Bayesian parameterization of photosynthesis models. *New Phytologist,* 200 (4), pp. 1132-1144.

Foley, J.A., Costa, M.H., Delire, C., Ramankutty, N. and Snyder, P. (2003) Green surprise? How terrestrial ecosystems could affect earth's climate. *Frontiers in Ecology and the Environment,* 1 (1), pp. 38-44.

Funk, J.L., Larson, J.E., Ames, G.M., Butterfield, B.J., Cavender-Bares, J., Firn, J., Laughlin, D.C., Sutton-Grier, A.,E., Williams, L. and Wright, J. (2017) Revisiting the Holy Grail: using plant functional traits to understand ecological processes. *Biological Reviews,* 92 (2), pp. 1156-1173.

Gates, D.M., Keegan, H.J., Schleter, J.C. and Weidner, V.R. (1965) Spectral properties of plants. *Applied Optics,* 4 (1), pp. 11-20.

Gökkaya, K., Thomas, V., Noland, T.L., McCaughey, H., Morrison, I. and Treitz, P. (2015) Prediction of macronutrients at the canopy level using spaceborne imaging spectroscopy and LiDAR data in a mixedwood boreal forest. *Remote Sensing,* 7 (7), pp. 9045-9069.

Hoffmann, W.A., Franco, A.C., Moreira, M.Z. and Haridasan, M. (2005) Specific leaf area explains differences in leaf traits between congeneric savanna and forest trees. *Functional Ecology,* 19 (6), pp. 932-940.

Houborg, R., Soegaard, H. and Boegh, E. (2007) Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data. *Remote Sensing of Environment,* 106 (1), pp. 39-58.

Houghton, R.A., House, J.I., Pongratz, J., Van Der Werf, G.,R., Defries, R.S., Hansen, M.C., Quéré, C.L. and Ramankutty, N. (2012) Carbon emissions from land use and land-cover change. *Biogeosciences,* 9 (12), pp. 5125-5142.

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F.W., Asner, G.P., Guralnick, R., Kattge, J., Latimer, A.M. and Moorcroft, P. (2016) Monitoring plant functional diversity from space. *Nature plants,* 2 (3), pp. 1-5.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Tautenhahn, S., Werner, G.D.A., Aakala, T. and Abedi, M. (2020) TRY plant trait database–enhanced coverage and open access. *Global Change Biology,* 26 (1), pp. 119-188.

Kattge, J., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P.B. and Wright, I.J. (2011) TRY–a global database of plant traits. *Global Change Biology,* 17 (9), pp. 2905-2935.

Knipling, E.B. (1970) Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment,* 1 (3), pp. 155-159.

Kokaly, R.F. (2001) Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment,* 75 (2), pp. 153-161.

Lavorel, S. and Garnier, E. (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Functional Ecology,* 16 (5), pp. 545-556.

Le Quéré, C., Peters, G.P., Andres, R.J., Andrew, R.M., Boden, T.A., Ciais, P., Friedlingstein, P., Houghton, R.A., Marland, G. and Moriarty, R. (2014) Global carbon budget 2013. *Earth System Science Data,* 6 (1), pp. 235-263.

Malhi, Y. (2002) Carbon in the atmosphere and terrestrial biosphere in the 21st century. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences,* 360 (1801), pp. 2925-2945.

McGill, B.J., Enquist, B.J., Weiher, E. and Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in ecology & evolution,* 21 (4), pp. 178-185.
McMahon, S.M., Harrison, S.P., Armbruster, W.S., Bartlein, P.J., Beale, C.M., Edwards, M.E., Kattge, J., Midgley, G., Morin, X. and Prentice, I.C. (2011) Improving assessment and modelling of climate change impacts on global terrestrial biodiversity. *Trends in Ecology & Evolution,* 26 (5), pp. 249-259.
Meir, P., Cox, P. and Grace, J. (2006) The influence of terrestrial ecosystems on climate. *Trends in Ecology & Evolution,* 21 (5), pp. 254-260.
Messier, J., McGill, B.J. and Lechowicz, M.J. (2010) How do traits vary across ecological scales? A case for trait-based ecology. *Ecology Letters,* 13 (7), pp. 838-848.
Metcalfe, D.B., Hermans, T.D.G., Ahlstrand, J., Becker, M., Berggren, M., Björk, R.,G., Björkman, M.,P., Blok, D., Chaudhary, N. and Chisholm, C. (2018) Patchy field sampling biases understanding of climate change impacts across the Arctic. *Nature ecology & evolution,* 2 (9), pp. 1443-1448.
Nunes, M.H., Both, S., Bongalov, B., Brelsford, C., Khoury, S., Burslem, D.F.R.P., Philipson, C., Majalap, N., Riutta, T. and Coomes, D.A. (2019) Changes in leaf functional traits of rainforest canopy trees associated with an El Niño event in Borneo. *Environmental Research Letters,* 14 (8), pp. 085005.
Pajares, G. (2015) Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs). *Photogrammetric Engineering & Remote Sensing,* 81 (4), pp. 281-330.
Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L. and Canadell, J.G. (2011) A large and persistent carbon sink in the world's forests. *Science,* 333 (6045), pp. 988-993.
Pearson, R.G., Phillips, S.J., Loranty, M.M., Beck, P.S.A., Damoulas, T., Knight, S.J. and Goetz, S.J. (2013) Shifts in Arctic vegetation and associated feedbacks under climate change. *Nature climate change,* 3 (7), pp. 673-677.
Poorter, H., Lambers, H. and Evans, J.R. (2014) Trait correlation networks: a whole-plant perspective on the recently criticized leaf economic spectrum. *New Phytologist,* 201 (2), pp. 378-382.
Poorter, L. and Bongers, F. (2006) Leaf traits are good predictors of plant performance across 53 rain forest species. *Ecology,* 87 (7), pp. 1733-1743.
Poorter, L. and Markesteijn, L. (2008) Seedling traits determine drought tolerance of tropical tree species. *Biotropica,* 40 (3), pp. 321-331.
Propastin, P., Ibrom, A., Knohl, A. and Erasmi, S. (2012) Effects of canopy photosynthesis saturation on the estimation of gross primary productivity from MODIS data in a tropical forest. *Remote Sensing of Environment,* 121, pp. 252-260.
Reich, P.B., Walters, M.B. and Ellsworth, D.S. (1997) From tropics to tundra: global convergence in plant functioning. *Proceedings of the National Academy of Sciences,* 94 (25), pp. 13730-13734.
Revill, A., Florence, A., MacArthur, A., Hoad, S., Rees, R. and Williams, M. (2020) Quantifying uncertainty and bridging the scaling gap in the retrieval of leaf area index by coupling Sentinel-2 and UAV observations. *Remote Sensing,* 12 (11), pp. 1843.
Rossatto, D.R., Hoffmann, W.A., Silva, L.d.C.R., Haridasan, M., Sternberg, L.S.L. and Franco, A.C. (2013) Seasonal variation in leaf traits between congeneric savanna and

forest trees in Central Brazil: implications for forest expansion into savanna. *Trees,* 27 (4), pp. 1139-1150.
Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M., Reeves, M. and Hashimoto, H. (2004) A continuous satellite-derived measure of global terrestrial primary production. *Bioscience,* 54 (6), pp. 547-560.
Sakschewski, B., von Bloh, W., Boit, A., Rammig, A., Kattge, J., Poorter, L., Peñuelas, J. and Thonicke, K. (2015) Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. *Global Change Biology,* 21 (7), pp. 2711-2725.
Santiago, L.S. (2007) Extending the leaf economics spectrum to decomposition: evidence from a tropical forest. *Ecology,* 88 (5), pp. 1126-1131.
Scheiter, S., Langan, L. and Higgins, S.I. (2013) Next-generation dynamic global vegetation models: learning from community ecology. *New Phytologist,* 198 (3), pp. 957-969.
Serbin, S.P., Dillaway, D.N., Kruger, E.L. and Townsend, P.A. (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *Journal of experimental botany,* 63 (1), pp. 489-502.
Serbin, S.P., Wu, J., Ely, K.S., Kruger, E.L., Townsend, P.A., Meng, R., Wolfe, B.T., Chlus, A., Wang, Z. and Rogers, A. (2019) From the Arctic to the tropics: multibiome prediction of leaf mass per area using leaf reflectance. *New Phytologist,* 224 (4), pp. 1557-1568.
Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J.O., Levis, S., Lucht, W. and Sykes, M.T. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology,* 9 (2), pp. 161-185.
Smith, M., Martin, M.E., Plourde, L. and 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 (6), pp. 1332-1337.
Stein, B.R., Thomas, V.A., Lorentz, L.J. and Strahm, B.D. (2014) Predicting macronutrient concentrations from loblolly pine leaf reflectance across local and regional scales. *GIScience & Remote Sensing,* 51 (3), pp. 269-287.
Suding, K.N. and Goldstein, L.J. (2008) Testing the Holy Grail framework: using functional traits to predict ecosystem change. *New Phytologist,* 180 (3), pp. 559-562.
Swinfield, T., Both, S., Riutta, T., Bongalov, B., Elias, D., Majalap-Lee, N., Ostle, N., Svátek, M., Kvasnica, J. and Milodowski, D. (2020) Imaging spectroscopy reveals the effects of topography and logging on the leaf chemistry of tropical forest canopy trees. *Global Change Biology,* 26 (2), pp. 989-1002.
Wagner, F.H., Hérault, B., Rossi, V., Hilker, T., Maeda, E.E., Sanchez, A., Lyapustin, A.I., Galvão, L.S., Wang, Y. and Aragão, L.,E.O.C. (2017) Climate drivers of the Amazon forest greening. *PloS one,* 12 (7), pp. e0180932.
Wang, A. and Price, D.T. (2007) Estimating global distribution of boreal, temperate, and tropical tree plant functional types using clustering techniques. *Journal of Geophysical Research: Biogeosciences,* 112.
Westoby, M. and Wright, I.J. (2006) Land-plant ecology on the basis of functional traits. *Trends in ecology & evolution,* 21 (5), pp. 261-268.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C. and Diemer, M. (2004) The worldwide leaf economics spectrum. *Nature,* 428 (6985), pp. 821-827.
Wright, I.J., Reich, P.B., Cornelissen, J.H.C., Falster, D.S., Groom, P.K., Hikosaka, K., Lee, W., Lusk, C.H., Niinemets, Ü. and Oleksyn, J. (2005) Modulation of leaf economic traits and trait relationships by climate. *Global Ecology and Biogeography,* 14 (5), pp. 411-421.
Wright, J.P. and Sutton-Grier, A. (2012) Does the leaf economic spectrum hold within local species pools across varying environmental conditions? *Functional Ecology,* 26 (6), pp. 1390-1398.
Wu, C., Wang, L., Niu, Z., Gao, S. and Wu, M. (2010) Nondestructive estimation of canopy chlorophyll content using Hyperion and Landsat/TM images. *International Journal of Remote Sensing,* 31 (8), pp. 2159-2167.

Chapter 2:

History and context



Svalbard
Photo: Eleanor Thomson

2.1. Airborne mapping of biochemical traits

2.1.1 Manned aircraft mapping of biochemical traits

Using a similar method to leaf spectroscopy, airborne imaging spectrometers measure solar radiation reflected from the vegetation canopy in narrow, contiguous spectral bands, covering a broad wavelength such as 350–2,500 nm. AVIRIS (Airborne Visible-Infrared Imaging Spectroscopy) was the first environmental imaging spectrometer of this type, mounted on a NASA ER-2 aircraft and flown in 1986 to study the environmental impacts of Southern California wildfires. Since then, AVIRIS has been used for a wide-variety of ecological applications (Singh *et al*. 2015; Huesca Martinez *et al*. 2019; Wang *et al*. 2019), but it does not have a specific traits or tropical forest focus. In contrast, the Global Airborne Observatory (formerly the Carnegie Airborne Observatory: CAO) combines an imaging spectrometer and light detection and ranging (LiDAR) technology onboard a Dornier 228-202 to quantify biochemical properties and vegetation structure, primarily across the tropics (Asner *et al*. 2007; 2012b). In 2016, Asner *et al*. produced the first maps of LMA, N and P over the entire Peruvian Amazon. These maps were followed in 2017, by maps of leaf calcium, phenols, lignin and water content over the same area. Similar work has been carried out in Hawaii and Borneo, using the CAO and replica manned platforms (Asner *et al.* 2008a, 2008b, 2016; 2017; Martin *et al.* 2018; Swinfield *et al*. 2020). These studies have revolutionized landscape-scale understanding of the spatial distribution and controls on leaf traits. For example, Asner *et al*. (2016) found that elevation and substrate were codominant drivers of leaf trait distributions, with topo-edaphic conditions mediating and at times eliminating classic LES relationships. Over Borneo, Swinfield *et al*. (2020) found that topography and forest disturbance were the main drivers of foliar chemistry, with N and P decreasing with increased elevation and forest disturbance.

2.1.2 UAV mapping of biochemical traits

In the last ten years, developments in robotics and computer vision, together with the miniaturization of sensing technologies, has opened up previously unavailable opportunities for UAV-based remote sensing (Colomina and Molina, 2014). Driven by

'smart farming' initiatives, UAVs, mounted with multispectral sensors designed to measure plant health and nutrient content, have become increasingly affordable and user friendly, based on 'out of the box' image collection and processing technology (Barbieri *et al*. 2018; Kim *et al*. 2019). Crop-based UAV studies have reported high accuracy values for monitoring a wide variety of agricultural crops for nitrogen, phosphorus and potassium deficiencies, water stress and disease (Zhu *et al*. 2009; Hoffmann *et al*. 2016; Severtson *et al*. 2016; Park *et al*. 2017; Hunt *et al*. 2018; Noguera *et al*. 2021). Expanding these studies to natural ecosystems, multispectral UAVs have been deployed to measure biochemical traits across grassland and Patagonian forests (Lu *et al*. 2018; Taylor-Zavala *et al*. 2021; Zhao *et al*. 2021). Yet, these kinds of studies are few and far between, with the majority of UAV-based ecological surveys directed at quantifying vegetation cover and species discrimination (Chen *et al*. 2016; Baena *et al*. 2017; Oldland *et al*. 2017; Cao *et al*. 2018; Melville *et al*. 2019; Théau *et al*. 2021).

2.1.3 Hyperspectral UAV mapping of biochemical traits

In contrast to multispectral cameras, hyperspectral imaging spectrometers provide numerous spectral bands that can capture fine-scale spectral differences and improve discrimination of vegetation properties. Hyperspectral UAV systems have been used to quantify biochemical traits in grasslands (Capolupo *et al*. 2015). Yet, across natural ecosystems, hyperspectral UAV studies are even rarer than multispectral UAV studies, likely due to the absence of affordable UAV-hyperspectral systems and user-friendly hypercube processing software. Following the trajectory of multispectral UAVs however, it is surely only a matter of time before the first low-cost hyperspectral system is developed, with user-friendly image processing software following.

2.1.4 Machine learning methods

In addition to the development of airborne platforms and spectral sensors, recent advances in machine learning methods have been essential to the remote quantification of leaf traits. Part of the fundamental innovation of machine learning methods is their efficiency in target-oriented learning of image features (traits) from raw input data (spectral imagery).

Many of these machine learning techniques are well-established mathematical techniques developed for predictive modelling, and have only recently been applied to spectral imagery. The strength of these methods lies in their ability to bypass the need for an *a priori* understanding of the relationship between leaf traits and optical imagery, and instead allow models to empirically 'learn' from field-measured trait values, using spectral bands as predictors. Questions have arisen concerning the application of these methods to spatial data which is inherently collinear, violating the assumption of observation independence shared by most machine learning procedures (Kühn, 2006; Ploton *et al.* 2020). Therefore, when applying machine learning methods to spatial data, it is important to consider issues of spatial autocorrelation and minimize these are far as possible. In particular, it is important to ensure that spatial observations are as independent as possible (e.g. not clustered) and, where necessary, to validate results using a spatially structured validation sample in order to avoid overly optimistic results generated as a product of spatial autocorrelation (Ploton *et al*. 2020).

2.1.5 Partial Least Squares Regression

In particular, partial least squares regression (PLSR) has emerged as powerful tool for leaf trait prediction (Haarland and Thomas, 1988). The PLSR technique is useful because it does not focus on a small number of wavelengths but uses the entire leaf spectral continuum in its analysis. PLSR is designed to handle situations where the number of predictor variables (e.g., 2150 spectral bands) is higher than the number of observations, avoiding earlier problems of collinearity and statistical overfitting (Wold *et al.* 1984). PLSR is particularly effective for spectral imagery, as it uses the continuous spectrum as a single measurement, rather than carrying out a band-by-band analysis.

The goal of PLSR is to predict Y (where Y is an n x 1 vector of trait predictions for each observation) from X (an n x m matrix of predictors for each observation) and to describe the common structure underlying the two variables. PLSR searches for a small set of components (latent variables) that simultaneously explain as much as the covariance in X, and as much as the covariance between X and Y, as possible (Haaland and Thomas, 1988). It is followed by a regression step, where the decomposition of X is used to predict Y. PLSR returns a linear model, where Y is predicted using a vector (m x 1) of regression

coefficients applied to X. These regression coefficients also inform the relative importance of each predictor to determining Y. PLSR models are generally using evaluated using k-fold cross-validation. K=3, is a classic evaluation technique, where 70% of the data is randomly selected for calibration of the models, with 30% reserved as an 'unseen' validation sample.

In addition to the strengths above, the advantages of PLSR are that it produces a linear model that can predict outside its calibration sample. However, PLSR also comes with the natural drawbacks of linear regression, in that the model cannot account for interactions between predictors, or relationships that are not linear. It also requires predictors measured in the same units, or at least standardized, and is sensitive to the relative scaling of the predictors (Cramer, 1993).

2.1.6 Random Forest Analysis

In the ecological literature, Random Forest has emerged a flexible, non-parametric machine learning method, which is particularly strong in the realm of classification, but can also be used for regression analysis (Breiman, 2001). Random Forest is often used for land cover classification and species classification (Immitzer *et al.* 2012; Rodriguez Galiano *et al*. 2012; Adelabu *et al*. 2015; Kulkani *et al*. 2016).

Random Forest is method for constructing predictive models, by 'growing' an ensemble of decision trees (Brieman, 2001). Random forest is based on the premise that a group of classifiers perform more accurately than a single classifier (Kotsiantis and Pintelas, 2004). Like PLSR, Random Forest is robust to high predictor collinearity, and situations where the number of predictors outnumber observations. In Random Forest analysis, each decision tree is constructed using a different bootstrap sample of the dataset ('bagging'). The data not used to grow the tree are reserved in an 'out-of-bag' (oob) subset to evaluate the performance of the tree. For each value of Y, all the decision trees which did not include that observation in their training data make a prediction. In random forest regression, each observation is assigned the average of all the predicted values. In random forest classification, each decision tree predicts a class, and the observation is assigned a class by the majority 'vote'. Model performance is measured by comparing predictions with the actual data. Predicter

importance is evaluated by permuting each predictor in turn and recording the loss in model accuracy.

The strengths of a random forest model are that it is a flexible, non-parametric machine learning classifier, which can account for a wide variety of non-linear relationships and interactions between predictors. Random forest is robust to a wide variety of predictors forming X, measured in different units. Its major weakness is that, for regression, it cannot predict outside the range of its training sample, and is even limited to predicting the exact values present in the training data.

2.1.7 Neural Networks

Recently, deep learning techniques, such as Convolutional Neural Networks (CNNs), have achieved prominence due to their high predictive performance for image classification (Albawi *et al*. 2017). For reliable results however, neural networks require extremely large datasets, often comprising thousands of individual measurements, which often exceeds the collection of trait campaigns. Thus, they are not included in this thesis.

2.1.8 UAV foliar trait mapping methods

Arguable, the main reason limiting the number of UAV-based trait mapping studies is not due to the lack of equipment access or scientific interest, but the absence of established methods linking field-sampled traits to UAV imagery (Houborg *et al*. 2015). While quantifying leaf traits using handheld spectroscopy and machine learning techniques, such as PLSR, is a relatively well-established approach, few studies have developed robust methodologies to measure canopy traits using UAV imagery in the same way. Thus, the first aim of this thesis is to develop a comprehensive methodology for the spatial quantification of biochemical leaf traits; both for widely-available multispectral UAV systems and rarer hyperspectral systems.

2.2. Optical remote sensing of process-based traits

2.2.1 Photosynthetic capacity

Plant photosynthetic capacity is mainly determined by two leaf traits; $Vcmax_{25}$ and $Jmax_{25}$. $Vcmax_{25}$ represents light-saturated conditions at 25°C, where photosynthesis is limited by the maximum carboxylation rate of RuBisCo, while $Jmax_{25}$ represents light-limited conditions at 25°C, where photosynthesis is limited by electron transport. These traits are fundamental parameters in the Farquhar, von Caemmerer and Berry (FvCB) model of photosynthesis (Farquhar *et al*. 1980), which is used by most TBMs to estimate gross primary productivity (GPP). As photosynthetic assimilation of CO_2 across the terrestrial biosphere represents the largest flux in the global carbon cycle, model outcomes are particularly sensitive to $Vcmax_{25}$ and $Jmax_{25}$ values (Knorr and Heimann, 2001; Zaehle *et al*. 2005; Friend *et al*. 2010;). Parameter uncertainty over $Vcmax_{25}$ has been found to result in ~30 Pg C $year^{-1}$ variation in annual global ecosystem productivity estimates (Bonan *et al.* 2011). As anthropogenic sources add ~9 Pg C to the atmosphere every year (Boden *et al*. 2017), there is considerable need to constrain $Vcmax_{25}$ and $Jmax_{25}$ values in terrestrial biosphere models (TBMs) to produce robust projections of global change.

Despite the fact that $Vcmax_{25}$ and $Jmax_{25}$ are known to vary considerably in space and time, TBMs typically assign a single, fixed $Vcmax_{25}$ and $Jmax_{25}$ parameter value to each plant functional type (PFT). At the tropics, $Vcmax_{25}$ and $Jmax_{25}$ are generally partitioned into two values– one for tropical evergreen trees and one for tropical deciduous trees (Rogers, 2014). Such generalisation fails to capture the magnitude, regional variation, and seasonality of photosynthetic capacity across tropical forest ecosystems, with obvious implications for vegetation and climate modelling.

Unsurprisingly therefore, the predictive performance of TBMs has been found to improve when parameter values are allowed to vary. At FLUXNET sites, which record high resolution data on carbon and water fluxes, model performance improved when $Vcmax_{25}$ and $Jmax_{25}$ were allowed to vary interannually according to site-calibrated values (Groenendijk *et al*. 2011). Additionally, some TBMs have improved the representation of

$Vcmax_{25}$ by simulating leaf nitrogen as part of the model and specifying a linear relationship between $Vcmax_{25}$ and leaf N for each PFT (e.g. Woodward *et al*. 1995; Kattge *et al*. 2009). Finally, Mercado *et al*. (2011) demonstrated considerable improvements to model predictions of carbon fluxes in the Amazon when information on leaf P was included.

2.2.2 Optical remote sensing of photosynthetic capacity

Remote sensing would be an ideal method of parametrizing TBMs with spatially-explicit $Vcmax_{25}$ and $Jmax_{25}$ values. In addition to its role in the terrestrial carbon cycle, $Vcmax_{25}$ has been found to be indicative of drought (Xu and Baldocchi, 2003), insect and pathogen damage (Dungan *et al*. 2007) and air pollution (Ainsworth *et al*. 2014).

To date however, most remote sensing of plant traits has focused on physical and chemical traits, that are either directly expressed in spectra, or highly correlated with traits that are. Process-based traits have typically received less attention in the spectroscopy literature, as they are not directly expressed in spectra, and their relationships with physical traits are not well understood. For example, photosynthetic capacity has been found to scale linearly with N (Field and Mooney, 1986; Hirose and Werger, 1987), but the scaling factor differs among biomes, often in relation to leaf P (Kattge *et al*. 2009, Walker *et al*. 2014; Domingues *et al*. 2015). Correlations between $Vcmax_{25}$ and LMA have been found at local scales, but this relationship appears less prominent at the global level (Walker *et al*. 2014).

Nevertheless, the few studies that have attempted to use spectroscopy to quantify photosynthesis capacity have been promising. Highest prediction accuracies of $Vcmax_{25}$ and $Jmax_{25}$ (r^2 = 0.88-0.92) have been achieved in crop studies (Serbin *et al.* 2012; Ainsworth *et al*. 2014; Heckmann *et al*. 2017), with models found to be robust to short-term variation in $Vcmax_{25}$ and $Jmax_{25}$ across a growing season (Serbin *et al*. 2012; Ainsworth *et al*. 2014; Barnes *et al*. 2017). Across forest ecosystems, Doughty *et al*. (2011) obtained lower prediction accuracies of $Vcmax_{25}$ and $Jmax_{25}$ ($r^2 \leq 0.52$) for 11 tropical tree species, while Dechant *et al*. (2017), achieved higher estimates (r^2 = 0.64-70) for 37 temperate deciduous tree species. The best photosynthetic capacity results (r^2 = 0.90)

were reported by Wu *et al*. (2019) for tropical forest leaves measured across Panama and Brazil.

Scaling from the leaf to the canopy level, Serbin *et al*. (2015) used AVIRIS imagery to achieve very good prediction results (r^2 = 0.90) of $Vcmax_{25}$ for nine crop species. However, to the best of our knowledge, no airborne study has attempted to derive photosynthetic capacity values across a natural ecosystem. Thus, this is the second aim of this thesis - to expand a focus from mapping biochemical leaf traits to developing predictive models of photosynthetic capacity across the tropics.

2.3. Satellite-based mapping of photosynthetic capacity

In contrast to most airborne sensors, satellites obtain lower spatial and spectral resolution imagery but benefit from global (or near global) coverage, and high revisit rates. The three main providers of open-access, multispectral imagery are Landsat, MODIS and Sentinel-2. Launched in 1972, Landsat is the longest running moderate-resolution (30m) Earth imaging programme. Landsat imagery contains 7 multispectral bands for ecological mapping and has a revisit time of 8 days. Launched in 1999, MODIS measures 19 spectral bands between 400-2500nm. MODIS has a spatial resolution of 250m-1km and a revisit time of 1-2 days.

Using Landsat and MODIS imagery to compute vegetation indices and relating these indices to biochemical plant properties is a well-established method for mapping leaf traits (Cohen and Goward, 2004; Homolová *et al*. 2013). NDVI, GCI, MSAVI2 and EVI have been used to estimate leaf chlorophyll content, leaf nitrogen content and LMA from local to global scales (Turner *et al*. 1999; Houborg *et al*. 2007; Wu *et al*. 2010). Using 'greenness' as a proxy for photosynthesis, MODIS takes this method one step further and calculates a daily GPP and annual net primary productivity (NPP) product (Running *et al*. 2004). However, this product has been found to saturate over areas of high leaf area index, such as tropical forests, and does not provide insight into the underlying photosynthetic traits of vegetation (Turner *et al*. 2006; Propastin *et al*. 2012; Cleveland *et al*. 2015).

At the local scale, Landsat imagery has been used to estimate $Vcmax_{25}$ directly, through relationships with leaf chlorophyll and nitrogen concentrations (Croft *et al*. 2016; Houborg *et al*. 2013). Vegetation indices derived from Landsat and MODIS data have also been used to estimate photosynthetic capacity parameters in temperate and alpine environments (Jin *et al*. 2012; Gao *et al*. 2014). However, vegetation indices rely on simple statistical relationships and assume that the researcher has an *a priori* understanding of the link between the photosynthetic capacity and the corresponding variation in reflectance. In addition, studies are often limited to using a few spectral bands, despite increased spectral range and resolution shown to increase predictive performance (Asner *et al*. 2011). Due to the spatial mismatch between a diverse vegetation canopy

(typically sampled at the level of an individual plant) and imagery pixel size, field surveys are typically only used in the validation stage and do not contribute to model creation. Higher spatial resolution imagery is required in order to directly link field-sampled traits to satellite imagery for model calibration, and avoid the need for an *a priori* understanding between the trait in question and its corresponding reflectance properties.

Launched in 2015, Sentinel-2 from the European Space Agency's (ESA) Copernicus programme represents the highest resolution, openly-accessible satellite imagery. Sentinel-2 has a ground sampling distance of 10m in the VNIR and 20m in the SWIR (less than the diameter of many tropical forest tree crowns), a revisit tine of 5 days and 13 spectral bands (*Figure 2.1*). Using all applicable Sentinel-2 spectral bands, and machine learning models calibrated on field-measured traits, Sentinel-2 imagery has been used to map biochemical leaf traits across tropical and temperate forests (Gara *et al*. 2019; Ma *et al*. 2019; Aguirre-Gutiérrez *et al*. 2020). To the best of our knowledge however, this thesis is the first research that uses Sentinel-2 imagery to create pan-tropical maps of $Vcmax_{25}$ and $Jmax_{25}$.

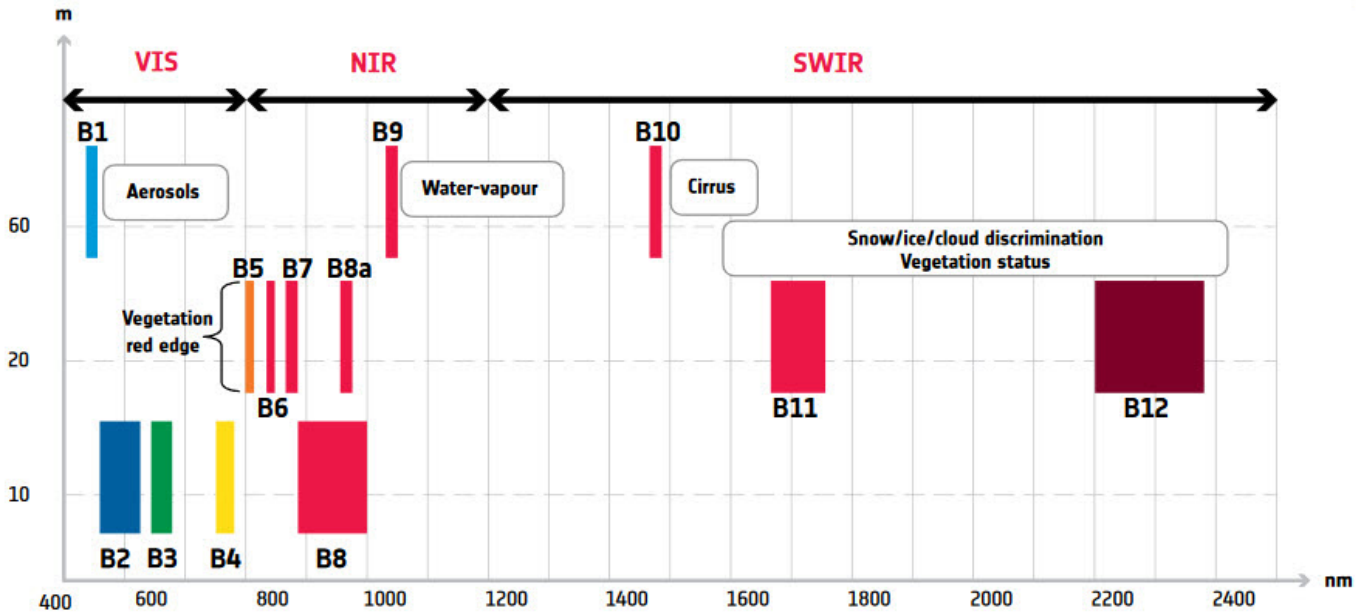


Figure 2.1: Spectral resolution (*x-axis*) against spatial resolution (*y-axis*) of Sentinel-2's 13 spectral bands. Source: https://www.nature.com/articles/s41598-021-83260-7/figures/3, accessed 15/11/2021.

2.4 References

Adelabu, S. and Dube, T. (2015) Employing ground and satellite-based QuickBird data and random forest to discriminate five tree species in a Southern African Woodland. *Geocarto International,* 30 (4), pp. 457-471.

Aguirre-Gutiérrez, J., Rifai, S., Shenkin, A., Oliveras, I., Bentley, L.P., Svátek, M., Girardin, C.A.J., Both, S., Riutta, T. and Berenguer, E. (2021) Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data. *Remote Sensing of Environment,* 252, pp. 112122.

Ainsworth, E.A., Serbin, S.P., Skoneczka, J.A. and Townsend, P.A. (2014) Using leaf optical properties to detect ozone effects on foliar biochemistry. *Photosynthesis Research,* 119 (1), pp. 65-76.

Albawi, S., Mohammed, T.A. and Al-Zawi, S. (2017) Understanding of a convolutional neural network. Ieee, pp. 1.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J.W. and Field, C.B. (2007) Carnegie airborne observatory: in-flight fusion of hyperspectral imaging and waveform light detection and ranging for three-dimensional studies of ecosystems. *Journal of Applied Remote Sensing,* 1 (1), pp. 013536.

Asner, G.P., Jones, M.O., Martin, R.E., Knapp, D.E. and Hughes, R.F. (2008a) Remote sensing of native and invasive species in Hawaiian forests. *Remote Sensing of Environment,* 112 (5), pp. 1912-1926.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J. and Hughes, R.F. (2008b) Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sensing of Environment,* 112 (5), pp. 1942-1955.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C., Carranza, L., Martinez, P., Houcheime, M., Sinca, F. and Weiss, P. (2011) Spectroscopy of canopy chemicals in humid tropical forests. *Remote Sensing of Environment,* 115 (12), pp. 3587-3598.

Asner, G.P., Knapp, D.E., Boardman, J., Green, R.O., Kennedy-Bowdoin, T., Eastwood, M., Martin, R.E., Anderson, C. and Field, C.B. (2012) Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. *Remote Sensing of Environment,* 124, pp. 454-465.

Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E. and Vaughn, N. (2016) Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proceedings of the National Academy of Sciences,* 113 (28), pp. E4043-E4051.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn, N.R. and Llactayo, W. (2017) Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science,* 355 (6323), pp. 385-389.

Baena, S., Moat, J., Whaley, O. and Boyd, D.S. (2017) Identifying species from the air: UAVs and the very high resolution challenge for plant conservation. *PloS one,* 12 (11), pp. e0188714.

Barnes, M.L., Breshears, D.D., Law, D.J., Van Leeuwen, W.,J.D., Monson, R.K., Fojtik, A.C., Barron-Gafford, G. and Moore, D.J.P. (2017) Beyond greenness: detecting

temporal changes in photosynthetic capacity with hyperspectral reflectance data. *PLoS One,* 12 (12), pp. e0189539.
Boden, T.A., Andres, R.J. and Marland, G. (2017) *Global, regional, and national fossil-fuel co2 emissions (1751-2014)(v. 2017).* Environmental System Science Data Infrastructure for a Virtual Ecosystem.
Bonan, G.B., Lawrence, P.J., Oleson, K.W., Levis, S., Jung, M., Reichstein, M., Lawrence, D.M. and Swenson, S.C. (2011) Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *Journal of Geophysical Research: Biogeosciences,* 116.
Breiman, L. (2001) Random forests. *Machine Learning,* 45 (1), pp. 5-32.
Cao, J., Leng, W., Liu, K., Liu, L., He, Z. and Zhu, Y. (2018) Object-based mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models. *Remote Sensing,* 10 (1), pp. 89.
Capolupo, A., Kooistra, L., Berendonk, C., Boccia, L. and Suomalainen, J. (2015) Estimating plant traits of grasslands from UAV-acquired hyperspectral images: a comparison of statistical approaches. *ISPRS International Journal of Geo-Information,* 4 (4), pp. 2792-2820.
Chen, J., Yi, S., Qin, Y. and Wang, X. (2016) Improving estimates of fractional vegetation cover based on UAV in alpine grassland on the Qinghai–Tibetan Plateau. *International Journal of Remote Sensing,* 37 (8), pp. 1922-1936.
Cleveland, C.C., Taylor, P., Chadwick, K.D., Dahlin, K., Doughty, C.E., Malhi, Y., Smith, W.K., Sullivan, B.W., Wieder, W.R. and Townsend, A.R. (2015) A comparison of plot-based satellite and Earth system model estimates of tropical forest net primary production. *Global Biogeochemical Cycles,* 29 (5), pp. 626-644.
Cohen, W.B. and Goward, S.N. (2004) Landsat's role in ecological applications of remote sensing. *Bioscience,* 54 (6), pp. 535-545.
Colomina, I. and Molina, P. (2014) Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of photogrammetry and remote sensing,* 92, pp. 79-97.
Cramer, R.D. (1993) Partial least squares (PLS): its strengths and limitations. *Perspectives in Drug Discovery and Design,* 1 (2), pp. 269-278.
Croft, H., Chen, J.M., Luo, X., Bartlett, P., Chen, B. and Staebler, R.M. (2017) Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global Change Biology,* 23 (9), pp. 3513-3524.
Dechant, B., Cuntz, M., Vohland, M., Schulz, E. and Doktor, D. (2017) Estimation of photosynthesis traits from leaf reflectance spectra: Correlation to nitrogen content as the dominant mechanism. *Remote Sensing of Environment,* 196, pp. 279-292.
Domingues, T.F., Ishida, F.Y., Feldpausch, T.R., Grace, J., Meir, P., Saiz, G., Sene, O., Schrodt, F., Sonké, B. and Taedoumg, H. (2015) Biome-specific effects of nitrogen and phosphorus on the photosynthetic characteristics of trees at a forest-savanna boundary in Cameroon. *Oecologia,* 178 (3), pp. 659-672.
Doughty, C.E., Asner, G.P. and Martin, R.E. (2011) Predicting tropical plant physiology from leaf and canopy spectroscopy. *Oecologia,* 165 (2), pp. 289-299.
Dungan, R.J., Turnbull, M.H. and Kelly, D. (2007) The carbon costs for host trees of a phloem-feeding herbivore. *Journal of Ecology,* 95 (4), pp. 603-613.

Farquhar, G.D., von Caemmerer, S.v. and Berry, J.A. (1980) A biochemical model of photosynthetic CO 2 assimilation in leaves of C 3 species. *Planta,* 149 (1), pp. 78-90.
Field, C.H. (1986) The photosynthesis-nitrogen relationship in wild plants. *On the economy of form and function,* .
Friend, A.D. (2010) Terrestrial plant production and climate change. *Journal of experimental botany,* 61 (5), pp. 1293-1309.
Gao, Y., Yu, G., Yan, H., Zhu, X., Li, S., Wang, Q., Zhang, J., Wang, Y., Li, Y. and Zhao, L. (2014) A MODIS-based Photosynthetic Capacity Model to estimate gross primary production in Northern China and the Tibetan Plateau. *Remote Sensing of Environment,* 148, pp. 108-118.
Gara, T.W., Darvishzadeh, R., Skidmore, A.K., Wang, T. and Heurich, M. (2019) Accurate modelling of canopy traits from seasonal Sentinel-2 imagery based on the vertical distribution of leaf traits. *ISPRS journal of photogrammetry and remote sensing,* 157, pp. 108-123.
Groenendijk, M., Dolman, A.J., Van der Molen, M., Leuning, R., Arneth, A., Delpierre, N., Gash, J., Lindroth, A., Richardson, A.D. and Verbeeck, H. (2011) Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data. *Agricultural and Forest Meteorology,* 151 (1), pp. 22-38.
Haaland, D.M. and Thomas, E.V. (1988) Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry,* 60 (11), pp. 1193-1202.
Heckmann, D., Schlüter, U. and Weber, A.P.M. (2017) Machine learning techniques for predicting crop photosynthetic capacity from leaf reflectance spectra. *Molecular plant,* 10 (6), pp. 878-890.
Hirose, T. and Werger, M.J.A. (1987) Nitrogen use efficiency in instantaneous and daily photosynthesis of leaves in the canopy of a Solidago altissima stand. *Physiologia Plantarum,* 70 (2), pp. 215-222.
Hoffmann, H., Jensen, R., Thomsen, A., Nieto, H., Rasmussen, J. and Friborg, T. (2016) Crop water stress maps for an entire growing season from visible and thermal UAV imagery. *Biogeosciences,* 13 (24), pp. 6545-6563.
Homolova, L., Malenovský, Z., Clevers, J.G.P.W., García-Santos, G. and Schaepman, M.E. (2013) Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity,* 15, pp. 1-16.
Houborg, R., Cescatti, A., Migliavacca, M. and Kustas, W.P. (2013) Satellite retrievals of leaf chlorophyll and photosynthetic capacity for improved modeling of GPP. *Agricultural and Forest Meteorology,* 177, pp. 10-23.
Houborg, R., Fisher, J.B. and Skidmore, A.K. (2015) *Advances in remote sensing of vegetation function and traits.* Elsevier.
Houborg, R., Soegaard, H. and Boegh, E. (2007) Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data. *Remote Sensing of Environment,* 106 (1), pp. 39-58.
Huesca Martinez, M., Zheng, T., Santos, M., Dattaraja, H.S., Suresh, H.S., Townsend, P.A., Sukumar, R. and Ustin, S. (2019) Woody plant diversity assessment and its

relationships to functional traits in Mundumalai National Park in India using AVIRIS-NG data. , pp. B22A.
Hunt, E.R., Horneck, D.A., Spinelli, C.B., Turner, R.W., Bruce, A.E., Gadler, D.J., Brungardt, J.J. and Hamm, P.B. (2018) Monitoring nitrogen status of potatoes using small unmanned aerial vehicles. *Precision agriculture,* 19 (2), pp. 314-333.
Immitzer, M., Atzberger, C. and Koukal, T. (2012) Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote sensing,* 4 (9), pp. 2661-2693.
Jin, P., Wang, Q., Iio, A. and Tenhunen, J. (2012) Retrieval of seasonal variation in photosynthetic capacity from multi-source vegetation indices. *Ecological Informatics,* 7 (1), pp. 7-18.
Kattge, J., Knorr, W., Raddatz, T. and Wirth, C. (2009) Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology,* 15 (4), pp. 976-991.
Kotsiantis, S. and Pintelas, P. (2004) Recent advances in clustering: A brief survey. *WSEAS Transactions on Information Science and Applications,* 1 (1), pp. 73-81.
Kim, J., Kim, S., Ju, C. and Son, H.I. (2019) Unmanned aerial vehicles in agriculture: A review of perspective of platform, control, and applications. *IEEE Access,* 7, pp. 105100-105115.
Knorr, W. and Heimann, M. (2001) Uncertainties in global terrestrial biosphere modeling: 1. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. *Global Biogeochemical Cycles,* 15 (1), pp. 207-225.
Kulkarni, A.D. and Lowe, B. (2016) Random forest algorithm for land cover classification.
Lu, B., He, Y. and Liu, H.H.T. (2018) Mapping vegetation biophysical and biochemical properties using unmanned aerial vehicles-acquired imagery. *International Journal of Remote Sensing,* 39 (15-16), pp. 5265-5287.
Ma, X., Mahecha, M.D., Migliavacca, M., van der Plas, F., Benavides, R., Ratcliffe, S., Kattge, J., Richter, R., Musavi, T. and Baeten, L. (2019) Inferring plant functional diversity from space: the potential of Sentinel-2. *Remote Sensing of Environment,* 233, pp. 111368.
Martin, R.E., Chadwick, K.D., Brodrick, P.G., Carranza-Jimenez, L., Vaughn, N.R. and Asner, G.P. (2018) An approach for foliar trait retrieval from airborne imaging spectroscopy of tropical forests. *Remote Sensing,* 10 (2), pp. 199.
Mercado, L., Quesada, C.A., Fyllas, N., Domingues, T., Anderson, L., Harris, P., Restrepo-Coupe, N. and Lloyd, J. (2011) Predicting terre firme forest stand level photosynthesis across the Amazon Basin.
Oldeland, J., Große-Stoltenberg, A., Naftal, L. and Strohbach, B.J. (2017) The potential of UAV derived image features for discriminating savannah tree species. In *The roles of remote sensing in nature conservation.* Springer, pp. 183-201.
Park, S., Ryu, D., Fuentes, S., Chung, H., Hernández-Montes, E. and O'Connell, M. (2017) Adaptive estimation of crop water stress in nectarine and peach orchards using high-resolution imagery from an unmanned aerial vehicle (UAV). *Remote Sensing,* 9 (8), pp. 828.

Propastin, P., Ibrom, A., Knohl, A. and Erasmi, S. (2012) Effects of canopy photosynthesis saturation on the estimation of gross primary productivity from MODIS data in a tropical forest. *Remote Sensing of Environment,* 121, pp. 252-260.
Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sanchez, J. (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing,* 67, pp. 93-104.
Rogers, A. (2014) The use and misuse of V c, max in Earth System Models. *Photosynthesis Research,* 119 (1), pp. 15-29.
Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M., Reeves, M. and Hashimoto, H. (2004) A continuous satellite-derived measure of global terrestrial primary production. *Bioscience,* 54 (6), pp. 547-560.
Serbin, S.P., Dillaway, D.N., Kruger, E.L. and Townsend, P.A. (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *Journal of experimental botany,* 63 (1), pp. 489-502.
Serbin, S.P., Singh, A., Desai, A.R., Dubois, S.G., Jablonski, A.D., Kingdon, C.C., Kruger, E.L. and Townsend, P.A. (2015) Remotely estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using imaging spectroscopy. *Remote Sensing of Environment,* 167, pp. 78-87.
Severtson, D., Callow, N., Flower, K., Neuhaus, A., Olejnik, M. and Nansen, C. (2016) Unmanned aerial vehicle canopy reflectance data detects potassium deficiency and green peach aphid susceptibility in canola. *Precision Agriculture,* 17 (6), pp. 659-677.
Singh, A., Serbin, S.P., McNeil, B.E., Kingdon, C.C. and Townsend, P.A. (2015) Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications,* 25 (8), pp. 2180-2197.
Swinfield, T., Both, S., Riutta, T., Bongalov, B., Elias, D., Majalap-Lee, N., Ostle, N., Svátek, M., Kvasnica, J. and Milodowski, D. (2020) Imaging spectroscopy reveals the effects of topography and logging on the leaf chemistry of tropical forest canopy trees. *Global Change Biology,* 26 (2), pp. 989-1002.
Taylor-Zavala, R., Ramírez-Rodríguez, O., de Armas-Ricard, M., Sanhueza, H., Higueras-Fredes, F. and Mattar, C. (2021) Quantifying Biochemical Traits over the Patagonian Sub-Antarctic Forests and Their Relation to Multispectral Vegetation Indices. *Remote Sensing,* 13 (21), pp. 4232.
Théau, J., Lauzier-Hudon, É., Aubé, L. and Devillers, N. (2021) Estimation of forage biomass and vegetation cover in grasslands using UAV imagery. *Plos one,* 16 (1), pp. e0245784.
Turner, D.P., Cohen, W.B., Kennedy, R.E., Fassnacht, K.S. and Briggs, J.M. (1999) Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment,* 70 (1), pp. 52-68.
Turner, D.P., Ritts, W.D., Cohen, W.B., Gower, S.T., Running, S.W., Zhao, M., Costa, M.H., Kirschbaum, A.A., Ham, J.M. and Saleska, S.R. (2006) Evaluation of MODIS NPP and GPP products across multiple biomes. *Remote Sensing of Environment,* 102 (3-4), pp. 282-292.
Walker, A.P., Beckerman, A.P., Gu, L., Kattge, J., Cernusak, L.A., Domingues, T.F., Scales, J.C., Wohlfahrt, G., Wullschleger, S.D. and Woodward, F.I. (2014) The relationship of leaf photosynthetic traits–Vcmax and Jmax–to leaf nitrogen, leaf phosphorus, and specific leaf area: a meta-analysis and modeling study. *Ecology and evolution,* 4 (16), pp. 3218-3235.

Wang, Z., Townsend, P.A., Schweiger, A.K., Couture, J.J., Singh, A., Hobbie, S.E. and Cavender-Bares, J. (2019) Mapping foliar functional traits and their uncertainties across three years in a grassland experiment. *Remote Sensing of Environment,* 221, pp. 405-416.
Wold, S., Ruhe, A., Wold, H. and Dunn, I., W.J. (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing,* 5 (3), pp. 735-743.
Woodward, F.I., Smith, T.M. and Emanuel, W.R. (1995) A global land primary productivity and phytogeography model. *Global Biogeochemical Cycles,* 9 (4), pp. 471-490.
Wu, C., Wang, L., Niu, Z., Gao, S. and Wu, M. (2010) Nondestructive estimation of canopy chlorophyll content using Hyperion and Landsat/TM images. *International Journal of Remote Sensing,* 31 (8), pp. 2159-2167.
Wu, J., Rogers, A., Albert, L.P., Ely, K., Prohaska, N., Wolfe, B.T., Oliveira Jr, R.C., Saleska, S.R. and Serbin, S.P. (2019) Leaf reflectance spectroscopy captures variation in carboxylation capacity across species, canopy environment and leaf age in lowland moist tropical forests. *New Phytologist,* 224 (2), pp. 663-674.
Xu, L. and Baldocchi, D.D. (2003) Seasonal trends in photosynthetic parameters and stomatal conductance of blue oak (Quercus douglasii) under prolonged summer drought and high temperature. *Tree physiology,* 23 (13), pp. 865-877.
Zaehle, S., Sitch, S., Smith, B. and Hatterman, F. (2005) Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics. *Global Biogeochemical Cycles,* 19 (3).
Zhao, Y., Sun, Y., Chen, W., Zhao, Y., Liu, X. and Bai, Y. (2021) The Potential of Mapping Grassland Plant Diversity with the Links among Spectral Diversity, Functional Trait Diversity, and Species Diversity. *Remote Sensing,* 13 (15), pp. 3034.
Zhu, J., Wang, K., Deng, J. and Harmon, T. (2009) Quantifying nitrogen status of rice using low altitude UAV-mounted system and object-oriented segmentation methodology, pp. 603.

Chapter 5:

Multiscale mapping of plant functional groups and plant traits in the High Arctic using field spectroscopy, UAV imagery and Sentinel-2 data.

Authors: Eleanor R. Thomson[1], Marcus P. Spiegel[1], Inge H.J. Althuizen[2], Polly Bass[3], Shuli Chen[4], Adam Chmurzynski[4], Aud H. Halbritter[2], Jonathan J. Henn[5], Ingibjörg S. Jónsdóttir[6], Kari Klanderud[7], Yaoqi Li[2,8], Brian S. Maitner[4], Sean T. Michaletz[9], Pekka Niittynen[10], Ruben E. Roos[7], Richard J. Telford[2], Brian J. Enquist[4], Vigdis Vandvik[2], Marc Macias-Fauria[1], Yadvinder Malhi[1]

Affiliations:

[1] *School of Geography and the Environment, University of Oxford, Oxford. UK.*

[2] *Department of Biological Sciences and Bjerknes Centre for Climate Research, University of Bergen, Bergen, Norway.*

[3]*Matanuska-Susitna College, University of Alaska, Anchorage, Alaska, United States.*

[4] *Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, United States.*

[5]*Department of Integrative Biology, University of Wisconsin-Madison, Madison, USA*

[6] *Faculty of Life and Environmental Sciences, University of Iceland, 102 Reykjavik, Iceland.*

[7]*Faculty of Environmental Sciences and Natural Resource Management, P.O. Box 5003, NO-1432 Aas, Norway.*

[8]*Institute of Ecology and Key Laboratory for Earth Surface Processes of the Ministry of Education, College of Urban and Environmental Sciences, Peking University, Beijing, China.*

[9]*Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, Canada.*

[10]*Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland.*

5.1. Chapter Statement

This chapter starts by using field spectroscopy to classify plant species and quantify biochemical plant traits (water content, nitrogen (N), phosphorus (P) and C:N ratios) in the High Arctic. It compares the prediction accuracy of field spectroscopy with the prediction accuracy of UAV-collected 5-band multispectral imagery for retrieving the same vegetation properties. It ends by evaluating two methods of upscaling a UAV map of plant water content to the surrounding landscape using Sentinel-2 imagery.

While it may seem odd to have an Arctic chapter in a thesis principally focusing on tropical forests, the High Arctic is an ideal environment for testing the primary objectives of this thesis. From a remote sensing perspective, the High Arctic is a simpler ecosystem than a tropical forest. Species richness is much lower and vegetation cover lacks the structure and confounding effects of a tropical forest canopy. Thus, the High Arctic is an ideal ecosystem in which to establish the baseline potential of multispectral UAV imagery for mapping biochemical leaf traits.

This chapter was published on 21st April, 2021 in *Environmental Research Letters* (https://doi.org/10.1088/1748-9326/abf464).

Results were presented at the Detecting High Latitude Ecosystem Change conference at the Scott Polar Institute in Cambridge, UK and the Svalbard Science Conference in Oslo, Norway. This work was an invited presentation at the Svalbard Integrated Arctic Earth Observing System (SIOS) Conference in St. Andrews, UK.

5.2. Author Contributions:

The ideas, research questions, study design, data processing, data analysis and writing presented in this chapter are the work of the first author, Eleanor Thomson. These activities greatly benefited from the contributions of Marcus Spiegel, Marc Macias-Fauria and Yadvinder Malhi.

Data collection was a team effort between Eleanor Thomson, Marcus Spiegel, Adam Chmurzynski, Yaoqi Li, Shuli Chen, Marc Macias Fauria and Yadvinder Malhi. Species identification, particularly moss identification, was carried out by Polly Bass and Pekka Niittynen. Chemical analysis was undertaken at the University of Arizona.

Vigdis Vandvik and Brian Enquist provided the logistical and financial support for the data collection, along with contributions from Inge Althuizen, Aud Halbritter, Jonathan Henn, Ingibjörg Jónsdóttir, Kari Klanderud, Brian Maitner, Sean Michaletz, Ruben Roos and Richard Telford.

All authors provided constructive comments during the writing of the manuscript.

5.3. Abstract

The Arctic is warming twice as fast as the rest of the planet, leading to rapid changes in species composition and plant functional trait variation. Landscape-level maps of vegetation composition and trait distributions are required to expand spatially-limited plot studies, overcome sampling biases associated with the most accessible research areas, and create baselines from which to monitor environmental change. Unmanned aerial vehicles (UAVs) have emerged as a low-cost method to generate high-resolution imagery and bridge the gap between fine-scale field studies and lower resolution satellite analyses. Here we used field spectroscopy data (400–2500 nm) and UAV multispectral imagery to test spectral methods of species identification and plant water and chemistry retrieval near Longyearbyen, Svalbard. Using the field spectroscopy data and Random Forest analysis, we were able to distinguish eight common High Arctic plant tundra species with 74% accuracy. Using partial least squares regression (PLSR), we were able to predict corresponding water, nitrogen, phosphorus and C:N values (r^2 = 0.61–0.88, RMSEmean = 12%–64%). We developed analogous models using UAV imagery (five bands: Blue, Green, Red, Red Edge and Near-Infrared) and scaled up the results across a 450 m long nutrient gradient located underneath a seabird colony. At the UAV level, we were able to map three plant functional groups (mosses, graminoids and dwarf shrubs) at 72% accuracy and generate maps of plant chemistry. Our maps show a clear marine-derived fertility gradient, mediated by geomorphology. We used the UAV results to explore two methods of upscaling plant water content to the wider landscape using Sentinel-2 imagery. Our results are pertinent for high resolution, low-cost mapping of the Arctic.

Key Words: Svalbard, Random Forest, PLSR, functional trait, tundra, moss, shrubs, remote sensing, bird cliff.

5.4. Introduction

The Arctic is the fastest warming region on earth (Box *et al.* 2019). Air temperatures have risen at twice the global rate (Stocker *et al.* 2013), driving changes to the structure and functioning of tundra ecosystems (Wahren *et al.* 2005; Joshua *et al.* 2016; Park *et al*. 2016; Oberbauer *et al.* 2013; Khorsand *et al*. 2015). Increased temperatures lead to shifts in vegetation cover and composition and accelerate belowground nutrient cycling and mineralization rates, with potentially important effects on global climate (Nadelhoffer *et al.* 1991; Sturm and Douglas, 2005; Higuera *et al.* 2008; Loranty *et al.* 2011; Pearson *et al.* 2013; Nauta *et al.* 2015; Davidson *et al.* 2016a).

Plant traits are a primary control on the distribution and functioning of plants and underpin vegetation–climate responses (e.g. Soudzilovskaia *et al.* 2013; Bjorkman *et al*. 2018; Mekonnen *et al*. 2018;). Traits related to the uptake and allocation of resources, such as leaf mass per area, leaf water content, leaf nitrogen content (N) and leaf phosphorus content (P) can affect growth rates, plant longevity, primary productivity, decomposition rates and biogeochemical cycling (Reich and Ellsworth, 1997; Wright *et al.* 2004; Cornwell *et al.* 2008; Díaz *et al*. 2016). Simultaneously, morphology-related traits, such as leaf area and plant height, can influence aboveground biomass, surface albedo and snow dynamics (Sturm and Douglas, 2005; Hudson and Henry, 2009; Juszak *et al.* 2014). Establishing spatial environmental–trait relationships could provide a quantitative basis to forecast the effects of climate change in the Arctic (Myers-Smith *et al.* 2019; Thomas *et al.* 2020). Yet observational plot-studies often fail to isolate environmental drivers or take place at a smaller scale than the community/landscape they represent. Hence, the biotic consequences of warming in the coldest environments remain poorly predicted (Metcalfe *et al.* 2018).

When variations in vegetation characteristics affect spectral reflectance, these changes can be measured using optical remote sensing techniques. In comparison to traditional field-based methods, using remote sensing to generate vegetation data is less laborious, less invasive, more cost-effective and spatially continuous. Using multispectral satellite-based sensors, the distribution of Arctic vegetation has been well-established at broad scales through datasets such as the Circumpolar Arctic Vegetation Map (CAVM: Raynolds

et al. 2019), which classifies 16 Arctic vegetation types at 1 km resolution. Using coarse-resolution satellite imagery, widespread vegetation change at high latitudes has also been well-documented, including pan-Arctic 'greening' and 'browning' trends (Myneni *et al.* 1997; Bhatt *et al.* 2013; Zhang *et al.* 2017; Myers-Smith *et al.* 2020), with important implications for albedo, active layer depth, permafrost dynamics, carbon cycling and wildlife (Blok *et al.* 2010, 2011; Loranty *et al.* 2011; Bonfils *et al.* 2012; Loranty and Goetz, 2012; Tape *et al.* 2016; Fauchald *et al.* 2017; Frost *et al.* 2018; Arndt *et al.* 2019; Ims *et al.* 2019). However, vegetation cover in the Arctic is extremely heterogeneous, varying at very fine spatial scales (Suvanto *et al.* 2014; Opedal *et al.* 2015; Räsänen and Virtanen, 2019), and observations do not always agree at the satellite and plot level (Guay *et al.* 2014; Pattison *et al.* 2015; Myers-Smith *et al.* 2020). Thus, while remote sensing offers enormous potential to generate large-scale vegetation data across the Arctic, such coarse spatial and spectral resolution is limited in its ability to capture the fine-scale dynamics of tundra plants or resolve the key drivers of observed vegetation trends (Assmann *et al.* 2020; Myers-Smith *et al.* 2020; Siewert and Olofsson, 2020).

In lower-latitude environments, high-resolution field spectroscopy (visible-near infrared (VNIR: 400–1100)) or visible-short-wave infrared (VSWIR: 400–2500 nm) is a well-established technique for gathering plant taxonomic and trait data with a high degree of accuracy (Asner *et al.* 2012; Serbin *et al.* 2014, 2019; Roth *et al.* 2015). Specific spectral reflectance and absorption features have been directly linked to concentrations of cellulose, lignin, chlorophyll, nutrients and water (see Curran *et al.* 1989 for a comprehensive synthesis). These plant traits can be combined to create unique chemical 'fingerprints' to distinguish individual genera or species (Asner *et al.* 2008; Asner and Martin, 2009; Durgante *et al.* 2013; Lopatin *et al.* 2017). Compared to multispectral remote sensing, field spectroscopy provides numerous spectral bands that can capture fine-scale spectral differences and improve discrimination of vegetation features. Spectroscopy studies are rarely extended to the tundra however, likely due to the challenging logistics and the fact that field spectrometers are not designed to accommodate the small size of tundra plant leaves. Spectroscopy studies carried out in the Arctic have mainly focused on classifying tundra communities (Buchhorn *et al.* 2013; Huemmrich *et al.* 2013; Bratsch *et al.* 2016; Davidson *et al.* 2016b; Beamish *et al.* 2017;), vegetation cover fraction (Liu *et al.* 2017), biophysical plant traits, such as vegetation

height and biomass (Buchhorn *et al.* 2013; Bratsch *et al.* 2017), and leaf chlorophyll content or 'greenness' (Buchhorn *et al.* 2013; Kennedy *et al.* 2020), but not individual species or plant chemical traits. Remote sensing of high-latitude vegetation would benefit from more spectroscopy studies, which can link plant characteristics to spectral features at specific wavelengths, as well as establish remote sensing baselines (i.e. if a vegetation characteristic cannot be determined using field spectroscopy, the characteristic is unlikely to be distinguished using coarser spatial and spectral resolution imagery).

By itself however, spectroscopy cannot generate the spatial and temporal data required to map and monitor biotic change in the Arctic. In order to bridge this 'scale-gap', unmanned aerial vehicles (UAVs) have emerged as platforms to upscale field data and provide nuance to satellite observations (Riihimäki *et al.* 2019; Assman *et al.* 2020; Siewert and Olofsson, 2020). Due to the miniaturization of technology and decreasing costs, UAV-mounted sensors can facilitate the mapping and monitoring of vegetation at previously unachievable spatial, spectral and temporal resolutions. Using hyper and multispectral imagery, UAVs have been used to map plant communities (Tømmervik *et al.* 2014; Mora *et al.* 2015), individual species (Baena *et al.* 2017; Lu *et al.* 2017; Oldeland *et al.* 2017), plant traits (Capolupo *et al.* 2015; Constantin *et al.* 2015; Thomson *et al.* 2018) and plant health (Lucieer *et al.* 2014; Baena *et al.* 2017; Dash *et al.* 2018; Turner *et al.* 2018). While some UAV studies have applied structure-from-motion methods to analyse biophysical plant characteristics in the Arctic (e.g. Fraser *et al.* 2016; Palace *et al.* 2018), or applied spectral techniques to classify high latitude plant communities (e.g. Tømmervik *et al.* 2014; Mora *et al*. 2015; Davidson *et al.* 2016b; Langford *et al.* 2016), to the best of our knowledge, this is the first study to use field spectroscopy and UAVs to identify high Arctic plant species and map vegetation chemistry across the tundra.

In this study, we investigate the use of field spectroscopy for predicting species identity and retrieving plant water and chemistry concentrations across a range of common species at three sites found near Longyearbyen, Svalbard. Based on the demonstrated feasibility of generating plant taxonomic and trait data using field spectra, we develop analogous models using UAV five-band multispectral imagery. Additionally, we explore methods of upscaling our UAV results to the wider landscape using Sentinel-2 imagery.

Specifically, we aim to:

(a) develop predictive models of species identity and biochemical plant traits (water, nitrogen, phosphorus and C:N concentrations) using field spectroscopy,

(b) develop predictive models of species identity and the same plant traits using UAV spectra,

(c) apply the UAV models to UAV imagery to derive spatially continuous maps of plant species and their corresponding trait values,

(d) explore two methods of upscaling the UAV models to the wider landscape using Sentinel-2 imagery.

5.5. Methods

Study sites

The data were collected at three High Arctic study sites around Longyearbyen on the Svalbard Archipelago, Norway (**Figure 5.1**). The sites comprised two contrasting environmental gradients and an experimental warming site. The sites are characterized by a dry Arctic climate with a mean average temperature of −2.6 °C and an average maximum and minimum monthly temperature of 7.6 °C and −11.4 °C. Precipitation is around 190 mm per year (2005–2018: Lawrimore *et al.* 2016). In all cases, underlying soils were typical cryosols with a thin organic layer on top of inorganic sediments (Jones *et al.* 2010).

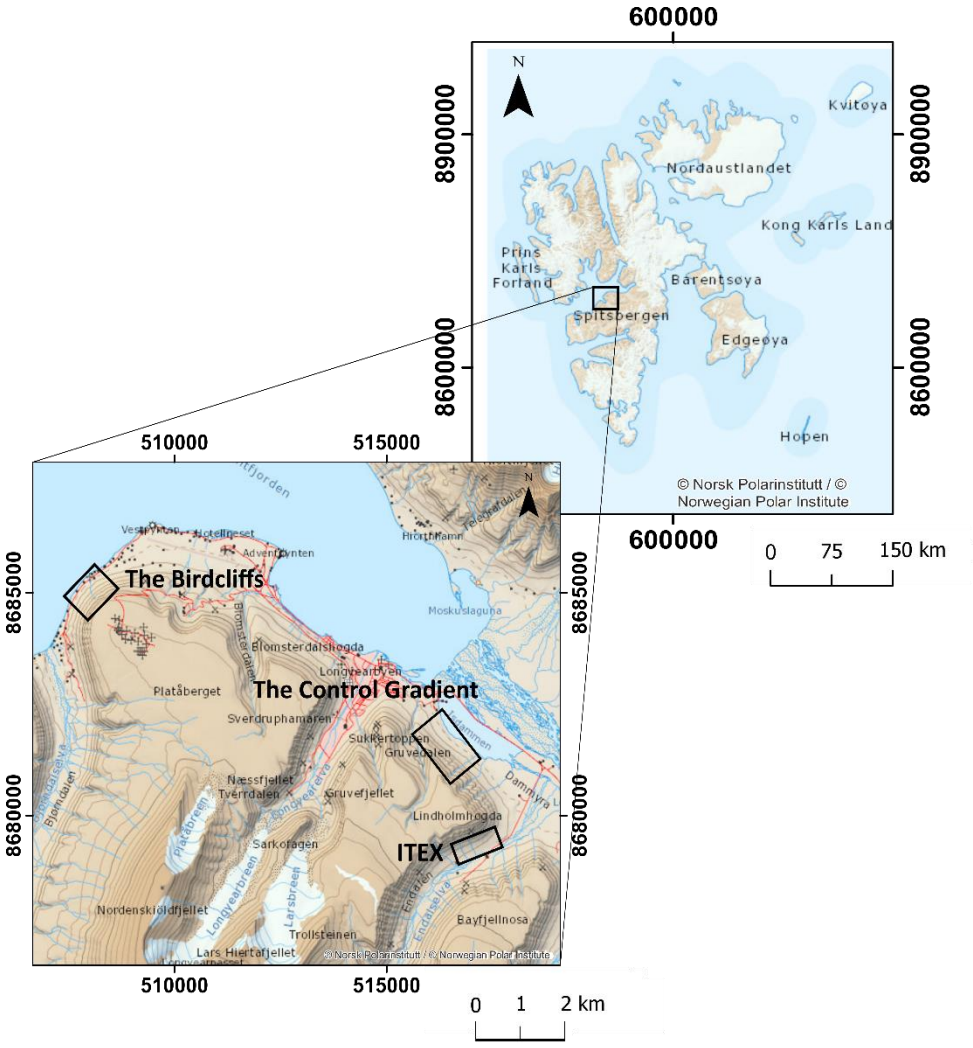


Figure 5.1: (a) Map of Svalbard archipelago. (b) Map of study sites near Longyearbyen, Svalbard. Basemaps are from the Norwegian Polar Institute. Coordinates in UTM Zone 33 N, WGS84 ellipsoid.

The Birdcliffs (78.14° N, 15.2° E)

Located near Bjørndalen, up the slope of Platåfjellet, the Birdcliffs is a 175 m high, 450 m long nutrient and elevation gradient, stretching from the base of a Little Auk (*Alle alle*) and Black-legged Kittiwake (*Rissa tridactyla*) colony to sea level at Isfjorden. The top of the gradient receives high nutrient input in the form of bird guano and other associated biological material (Zwolicki *et al.* 2013). The majority of the site is characterised by steep slopes and large alluvial fans made up of a mix unconsolidated material of various sizes. Vegetation is dominated by dry dwarf shrub tundra in topographically elevated areas, particularly *Salix polaris* and *Dryas octopetala*. Mosses and graminoids dominate in moist areas, especially around the edge of the alluvial fans and towards the flatter end of the gradient near the shore.

The Control Gradient (78.12° N, 15.4° E)

The Control Gradient is an 850 m long colluvial slope located up the north face of Lindholmhøgda, a spur at the mouth of the Adventdalen valley. The gradient runs from 200 m.a.s.l to a reservoir (Isdammen) at sea level. The area is characterised by dry dwarf shrub tundra that alternates with ridge communities of scarce vegetation. Habitats with thin winter snow cover are dominated by *Dryas octopetala* and *Salix polaris* (*Dryas* heath), while habitats of intermediate snow depth are dominated by Cassiope tetragona (*Cassiope* heath).

The International Tundra Experiment (78.18° N, 15.77° E)

The International Tundra Experiment (ITEX: Henry and Molau, 1997) is an experimental warming site located in Endalen. Data were collected outside the warming chambers from three predominant habitats: an exposed and relatively dry *Dryas* heath with shallow snow cover in winter, a mesic *Cassiope* heath with intermediate winter snow depth and a moist snowbed community dominated by bryophytes, *Salix polaris*, *Bistorta vivipara* and graminoids.

UAV imagery acquisition

All fieldwork was conducted as part of the Plant Functional Trait Course 4 (https://plantfunctionaltraitscourses.w.uib.no/), from 16th July 2018 to 27th July 2018. See **Figure 5.2** for a diagrammatic representation of the data collection and analysis.

UAV imaging data were acquired from all three sites. The data were acquired using a 3DR Solo drone equipped with a MicaSense RedEdge-MX multispectral camera and MicaSense RedEdge Downwelling Light Sensor (DLS). The MicaSense RedEdge–MX camera captures surface reflectance at five narrow spectral bands: blue (475 nm), green (560 nm), red (668 nm), Red Edge (717 nm) and NIR (840 nm). The DLS points upwards and captures illumination conditions, which are embedded in the metadata of each image. The drone was flown at an altitude of 40–60 m, resulting in <10 cm resolution imagery. To map the entirety of the study sites, multiple overlapping flights were required, which were taken over the course of one day at each site. Radiometric calibration images were recorded using a MicaSense reflectance panel as the calibration target. Ground control points for georeferencing were taken using the Emlid Reach+ differential GNSS system (Emlid, Hong Kong).

The imagery was processed in Pix4Dmapper (v.4.3.31, Pix4D, Lausanne, Switzerland) using a standard structure-from-motion (SfM) workflow. For each site, images from all flights were processed in the same project to form a single orthomosaic. The GCPs were manually identified in the images and georeferenced using their field-collected RTK–GNSS coordinates. Across the three sites, this workflow resulted in RMS errors of 0.10–0.14 m. Radiometric calibration was performed in Pix4D using the MicaSense reflectance target images and the metadata from the DLS.

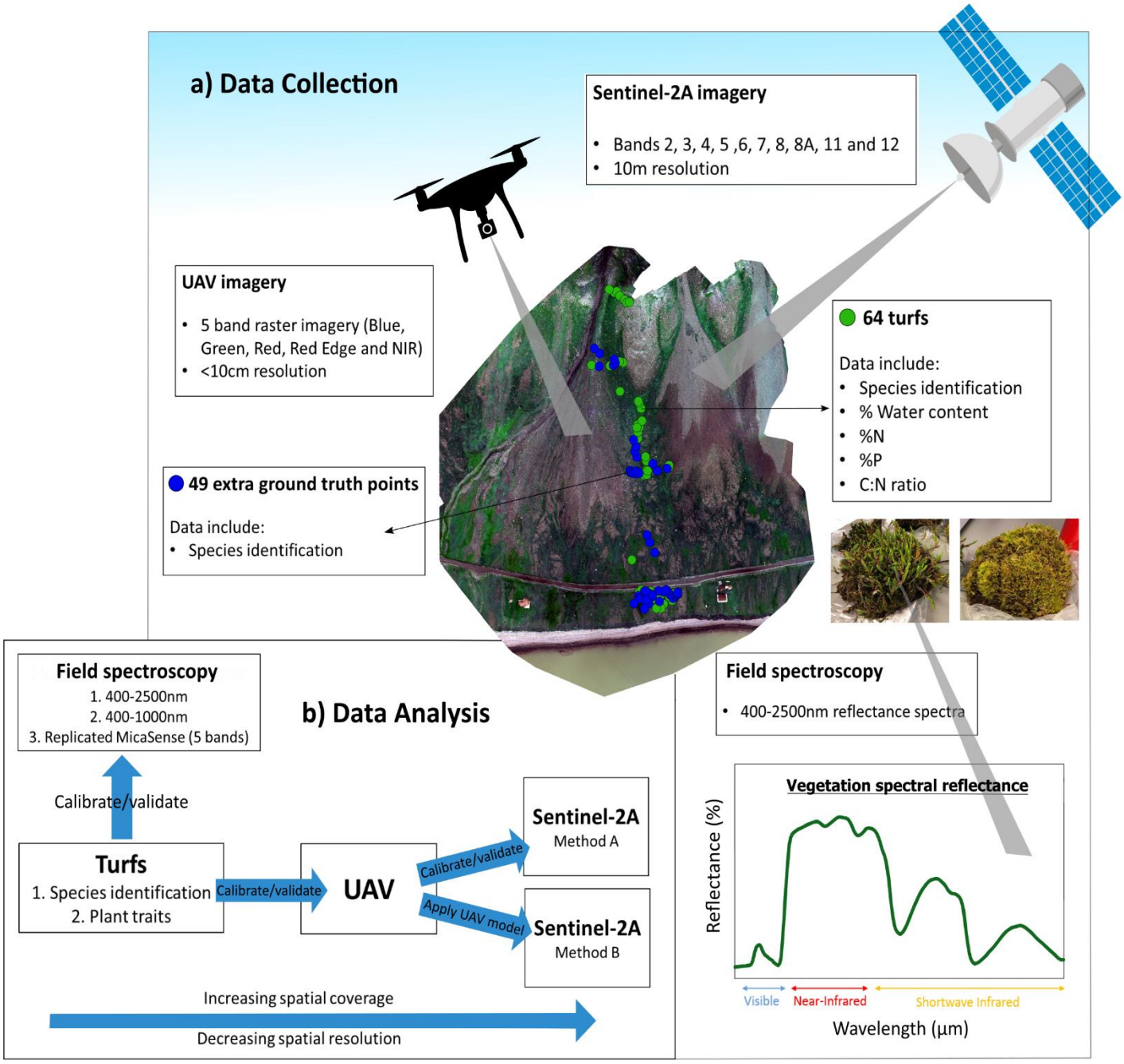


Figure 5.2: Diagrammatic representation of (a) the data collection and (b) the data analysis. Only the Birdcliffs site is shown here. Further turfs, ground-truth points and UAV imagery were collected at the Control Gradient and ITEX site.

Vegetation sampling

After the UAV imagery acquisition, 68 20 × 20 cm single-species turfs were collected from across the Birdcliffs and ITEX sites. The turfs were selected to represent the most common plant functional types identified across all sites: mosses, graminoids and dwarf shrubs (**Table 5.1, Figure 5.2**). High-accuracy GNSS coordinates were taken from the locations of the extracted turfs with an Emlid Reach+ differential GNSS system. Additional vegetation ground-truthing points were taken from all three sites (**Table 5.2, Figure 5.2**). The turfs were cut to a substrate depth of approximately 5 cm, sealed inside plastic bags and transported back to the University Centre in Svalbard (UNIS) for species identification and analysis.

Some shrub samples collected from the dry and mesic heaths at ITEX showed signs of tissue degradation, probably due to drought or frost damage (Bjerke *et al*. 2017). These samples were included in the plant trait analyses, as part of a continuum of trait values, but excluded from the species classification analyses. Mapping plant health falls outside the scope of this study and only healthy shrub communities were observed over the UAV upscaling areas.

Field spectroscopy measurements

Field spectroscopy measurements of the turf samples (350–2500 nm) were taken at UNIS using an ASD Fieldspec Pro with fibre optic cable and contact probe (Analytical Spectral Devices, Boulder, CO, USA). Turfs were stored outside and reflectance measurements were taken within 24 h of turf cutting. Spectroscopy measurements were not taken '*in situ*' at study sites, due to the non-portable set-up of the spectrometer. Turfs were selected to be as homogenous as possible. If multiple plant species were present across the turf, measurements were only taken from areas where the main species dominated. The contact probe was pushed firmly down onto the turf so all extraneous light was excluded from the measurement. Five measurements were taken at different locations across each turf. Each measurement consisted of 40 internally averaged reflectance readings to increase the signal-to-noise ratio. The spectrometer was optimised and calibrated for dark current and white light after every turf. For all statistical analyses, the

spectral data were trimmed to the 400–2500 nm range. For some analyses, the five measurements were averaged to form one spectrum per turf.

Plant trait sampling

After the field spectroscopy measurements had been taken, three 5 × 5 cm vegetation samples were cut from each turf. For each sample, all vegetation above the substrate was harvested. The samples were weighed for fresh mass, dried at 60 °C for 48 h and re-weighed for dry mass. Vegetation water content was calculated as ((fresh mass − dry mass)/fresh mass) × 100.

Nutrient analysis of the dried samples was carried out at the University of Arizona. The protocol to obtain total phosphorus concentration involved using persulfate oxidation followed by the acid molybdate method (APHA, 1992), after which P concentration was determined via colorimetric analysis with a spectrophotometer (ThermoScientific Genesys20, United States). Carbon, nitrogen, and their stable isotope ratios were measured at the Department of Geosciences Environmental Isotope Laboratory on a continuous-flow gas-ratio mass spectrometer.

The three samples were averaged to form one trait value per turf. See **Table S5.1** (available online at stacks.iop.org/ERL/16/055006/mmedia) for average trait values.

Table 5.1: Information on turf samples, including site location, plant functional type, genus and species. Turfs measured 20×20 cm. Five independent spectrum readings were taken per turf.

Turf samples					
Site	**Plant Functional Type**	**Genus**	**Species**	***N. of turf samples***	***N. of Spectra***
Birdcliffs	Moss	*Aulacomnium*	*palustre*	3	15
		Aulacomnium	*turgidum*	4	20
		Polytrichastrum	*alpinum*	2	10
		Polytrichum	*hyperboreum*	1	5
		Polytrichum	*strictum*	1	5
		Racomitrium	*canescens*	2	10
		Racomitrium	*lanuginosum*	2	10
Birdcliffs	Graminoid	*Alopecurus*	*ovatus*	7	35
		Luzula	*confusa*	6	30
		Luzula	*nivalis*	1	5
Birdcliffs	Dwarf shrub	*Cassiope*	*tetragona*	2	10
		Dryas	*octopetala*	5	25
		Salix	*polaris*	9	45
ITEX	Moss	*Identity not recorded*		5	25
ITEX	Graminoid	*Identity not recorded*		1	5
ITEX	Dwarf shrub	*Cassiope*	*tetragona*	2	10
	Dwarf Shrub	*Dryas*	*octopetala*	3	15
	Dwarf Shrub	*Salix*	*polaris*	3	15
	Senesced Shrub	*Cassiope*	*tetragona*	6	30
	Senesced Shrub	*Dryas*	*octopetala*	3	15
				Total: 68	**Total: 340**

Table 5.2: Information on ground-truthing vegetation survey, including site, plant functional type and number of data points. GNSS locations were taken with an Emlid Reach+ differential GNSS.

Ground-truthing vegetation survey		
Site	**Plant Functional Type**	***N. data points***
Birdcliffs	Moss	15
	Graminoid	23
	Dwarf Shrub	24
Control Gradient	Moss	9
	Graminoid	0
	Dwarf shrub	9
ITEX	Moss	6
	Graminoid	1
	Dwarf shrub	30
		Total: 117

5.6. Field spectroscopy analyses

Field spectroscopy-species analysis

The turf spectra were divided into eight vegetation classes representing a mixture of families, genera and species. Some of the moss and graminoid species were combined into their families or genera to increase the class sample size and because they were assumed to be spectrally similar. Mosses were classified into *Aulacomnium spp*. (containing *Aulacomnium palustre* and *Aulacomnium turgidum*); *Polytrichaceae spp*. (containing *Polytrichum hyperboreum, Polytrichastrum alpinum* and *Polytrichastrum strictum*); and *Racomitrium spp.* (containing *Racomitrium canescens* and *Racomitrium lanuginosum*). The graminoids were classified into *Alopecurus ovatus* (a grass) and *Luzula spp*. (a rush, containing *Luzula confusa* and *Luzula nivalis*). The dwarf shrubs were separated into *Cassiope tetragona, Dryas octopetala* and *Salix polaris*.

To quantify similarity or differences between the spectra, agglomerative hierarchical cluster analysis was performed using the eight classes described above and one averaged spectrum per turf. Cluster analysis works by calculating the distance (difference)

between each possible pair of spectra. The two spectra closest to each other are merged and the process repeats, with the number of clusters reduced by one each cycle. Spectral distances were calculated using Euclidean Distance in MATLAB (MathWorks, Natick, MA, USA).

The eight vegetation classes were predicted using five spectra per turf and Random Forest classification (Breiman , 2001). The five spectra per turf were used as independent samples, as the spectral curves were highly diverse due to the heterogeneous nature of the turfs. Random Forest is a machine learning classifier that grows as an ensemble of decision trees (Kotsiantis and Pintelas, 2004; Ghimire *et al.* 2010). Random Forest is useful for classifying hyperspectral data as it is robust against overfitting and can be used when the number of predictor variables (e.g. 2150 wavelength bands) is greater than the number of samples (e.g. 470 spectra). All the Random Forest analyses were carried out in MATLAB using the Treebagger function.

Field spectroscopy-trait analysis

Plant water, nitrogen, phosphorous and C:N values were predicted using one averaged spectrum per turf and partial least squares regression (PLSR: Haaland and Thomas, 1988). The PLSR method is effective, as it uses the continuous spectrum as a single measurement, rather than carrying out a band-by-band analysis and reduces a large predictor matrix (2150 spectral bands) down to a few relatively uncorrelated factors (known as latent variables). For the trait predictions, we chose to use PLSR instead of Random Forest Regression because, unlike Random Forest Regression, PLSR is able to extrapolate beyond its calibration data. While Random Forest is a strong tool for classification, it cannot predict values outside of its training range, limiting its accuracy when scaling up continuous data over wider areas that may contain trait values outside the range of the sample data.

PLSR analyses were carried out using the PLSregress function in MATLAB. The PLSR models were validated using an unseen 30% of the dataset. Due to the random nature of the 70:30 split, 1000 PLSR iterations were made for each spectra-trait analysis. The PLSR equations resulting in robust models (r^2 > the mean value of the 1000 iterations) were

evaluated using r^2 for the independent testing (val) dataset and RMSE as a percentage of the sample mean (as in Asner *et al.* 2015; Asner and Martin, 2015).

All Random Forest and PLSR analyses were carried out using (a) the full hyperspectral range from 400 to 2500 nm, (b) just the VIS–NIR region from 400 to 1000 nm and (c) downscaled spectra to match the five bands represented by the MicaSense camera (blue, green, red, red edge and near-infrared). The three analyses were carried out in order to act as a baseline from which the effects of decreased spectral information and the effects of using an airborne camera could be accurately separated.

UAV spectral analysis

UAV spectra-species analysis

UAV spectra were extracted from the orthomosaic at all the points where the turfs and ground-truthed species coordinates were taken. The spectra were extracted using the 'extract' function in the Raster package (Hijmans *et al.* 2015) in R (v.3.6.0, R Core Team, 2019). First, an NDVI mask was applied to the imagery to mask out any pixels with an NDVI of <0.5. The threshold of 0.5 was chosen as it represents the lower limit of NDVI values measured for vegetation in the Arctic (Bokhorst *et al.* 2012). The UAV spectra were extracted using three methods: simple (only the pixel value the coordinate falls in is returned), bilinear (the returned value is interpolated from the values of the four nearest pixels) and buffer (the returned value is averaged from all the pixels that fall within a specified 3 × 3 pixel buffer area). Across all analyses, the buffer method demonstrated the most robust results, hence only the results from these analyses are shown.

Using the UAV spectra, Random Forest classification was used to predict three broad plant functional groups: mosses, graminoids and dwarf shrubs. These plant functional groups represented the finest classification of vegetation that could be achieved using the UAV data and the Random Forest method of prediction. The Random Forest model that demonstrated the best result on the training/validation data was scaled up and applied to each pixel of the Birdcliff imagery to generate a continuous ecosystem map of plant functional type. The models were only applied to the Birdcliff site, as this was where the majority of the turfs were collected. As such, we only had confidence in the model validation for the species and environmental conditions found at this site.

UAV spectra-trait analysis

Trait values were predicted using the same UAV spectra and PLSR method described above. The PLSR model that maximised the r^2 and minimised the %RMSE was chosen and applied to each pixel of the UAV imagery to generate continuous maps of trait information. Trait values were also compared to an NDVI map generated from the UAV imagery to quantify the benefits of using a multispectral machine learning approach to trait prediction, compared to an NDVI-based prediction (**Figure S5.3**).

Sentinel-2 scaling analysis

Sentinel-2 imagery was used to explore methods of upscaling the UAV maps to the wider region. Plant water content was chosen as an example trait, as water has a strong direct expression in spectra and the Birdcliff site was assumed to be a nutrient anomaly in the landscape.

The Sentinel image used in the analysis was a Level-2A cloud-free, atmospherically-corrected image taken on 30th July 2018, one week after the field-collection. Only vegetation pixels were included in the analysis, identified using the Sentinel-2 scene classification layer (SCL). Analyses were carried out in R (v.3.6.0, R Core Team, 2019) and Google Earth Engine (https://earthengine.google.org/).

Two methods of upscaling from UAV to satellite data were tested:

(a) the UAV-generated map of plant water content (**Figure 5.8**) was resampled to match Sentinel-2 resolution (10 m) and gridlines using bilinear interpolation. Lower-resolution Sentinel bands were also resampled to 10 m using bilinear interpolation. The resampled UAV water values were used as a calibration and validation dataset for the Sentinel imagery. Thus, each Sentinel pixel (n = 329) had a single water value as well as a reflectance value for each relevant Sentinel band (Band 2 (492 nm), 3 (560 nm), 4 (665 nm), 5 (704 nm), 6 (741 nm), 7 (783 nm), 8 (833 nm), 8A (865 nm), 11 (1614 nm) and 12 (2202 nm)). Using this dataset and the same PLSR method described above, model coefficients were generated for each Sentinel band and applied to each pixel of the Sentinel image.

(b) the coefficients from the PLSR model used to create the UAV map of water content were applied to the Sentinel bands that best matched the MicaSense bands (Sentinel bands 2,3,4,5 and 8) for each pixel of the Sentinel image.

5.6. Results

Field spectroscopy identification of plant functional types and species

Average reflectance spectra for three plant functional types (moss, graminoids and dwarf shrubs) showed distinct spectral separability, especially in the NIR (750–1300 nm) and SWIR regions (1500–2500 nm, **Figure 5.3(a)**). Moss displayed the lowest reflectance across the NIR–SWIR region, especially in the water absorption bands near 1450, 1940 and 2500 nm (Curran *et al.* 1989). Within each plant functional type, unique spectral signatures were displayed by genera and species (**Figures 5.3(b)–(d)**). Shrubs were more spectrally similar to each other, but showed some divergence in the NIR–SWIR. Intraspecific spectral variability was low across most of the mosses, while *Alopecurus ovatus* and *Cassiope tetragona* displayed the highest spectral variability, especially across the NIR.

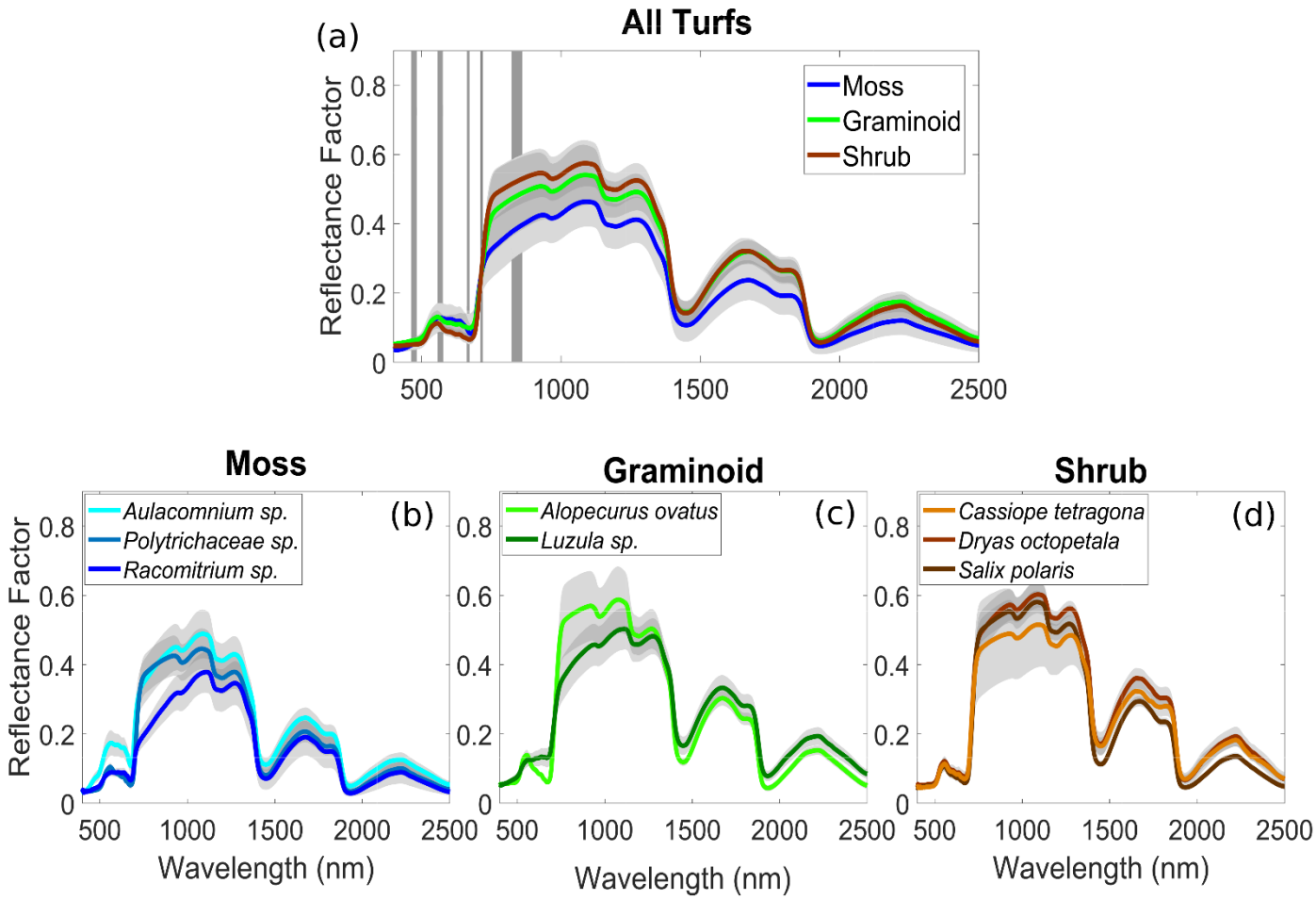


Figure 5.3: Mean (±standard deviation) of turf spectral reflectance (400–2500 nm). Reflectance was measured using an ASD Fieldspec Pro with fibre optic cable and contact probe. (a) Spectra were grouped into three plant functional types; mosses, graminoids and dwarf shrubs. The MicaSense RedEdge-MX five spectral bands (blue, green, red, red edge and near–infrared) are shown in grey for comparison purposes. (b) Mosses were classified into *Aulacomnium spp*. (*A. palustre* and *A turgidum*), *Polytrichaceae spp*. (*P. hyperboreum, P. alpinum* and *P. strictum*), and *Racomitrium spp.* (*R. canescens* and *R. lanuginosum*). (c) Graminoids were classified into *Alopecurus ovatus* and *Luzula spp*. (*L. confusa* and *L. nivalis*). (d) Shrubs were separated into *C. tetragona, D. octopetala* and *S. polaris.*

Using cluster analysis to further quantify spectral separability, mosses emerged as a distinct spectral family (**Figure 5.4**). All moss spectra clustered together, with some stratification of genera, especially *Aulacomnium* and *Racomitrium.* There was less guild clustering of graminoids and shrubs, however *Salix polaris* emerged as a tight cluster and most *Dryas octopetala* spectra clustered together. Despite some grouping of *Luzula*, the graminoids did not cluster together as a functional group. The majority of the spectral confusion was created by *Alopecurus ovatus* and *Cassiope tetragona,* whose high spectral variability meant they were commonly paired with other shrub and graminoid species.

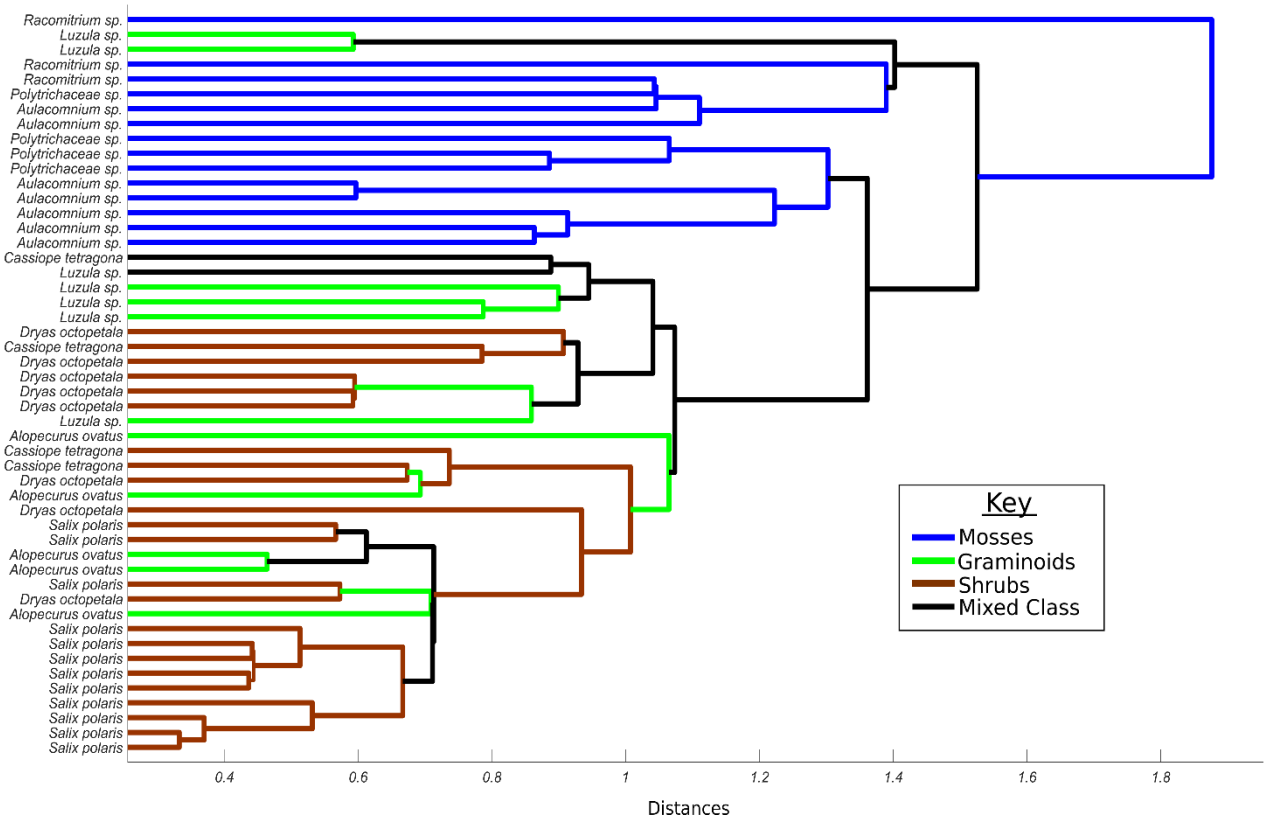


Figure 5.4: Results of spectral agglomerative hierarchical cluster analysis. One average spectrum per turf was used (n = 53). Spectra are labelled according to the eight vegetative classes described in **Figure 5.3** and coloured according to plant functional type (moss, graminoid, dwarf shrub).

Using Random Forest Classification and the full hyperspectral range (400–2500 nm), our results show that eight common Arctic tundra species can be classified with 74% accuracy (much above the expected statistical random accuracy value of 12.5%, given eight classes, **Figure 5.5**). Reflecting the results of the cluster analysis, moss species displayed the highest classification accuracy values (85, 80 and 80%), along with *Salix polaris* (80%) and *Luzula sp*. (80%). Reducing the amount of spectral information led to a small decrease in accuracy to 69 and 66% for the 400–1000 nm spectral range and downscaled MicaSense bands, respectively. Across the eight classes, the greatest source of uncertainty came from the classification of *Cassiope tetragona* and *Dryas octopetala*, which were commonly confused with each other (see **Figure S5.1** for confusion matrices).

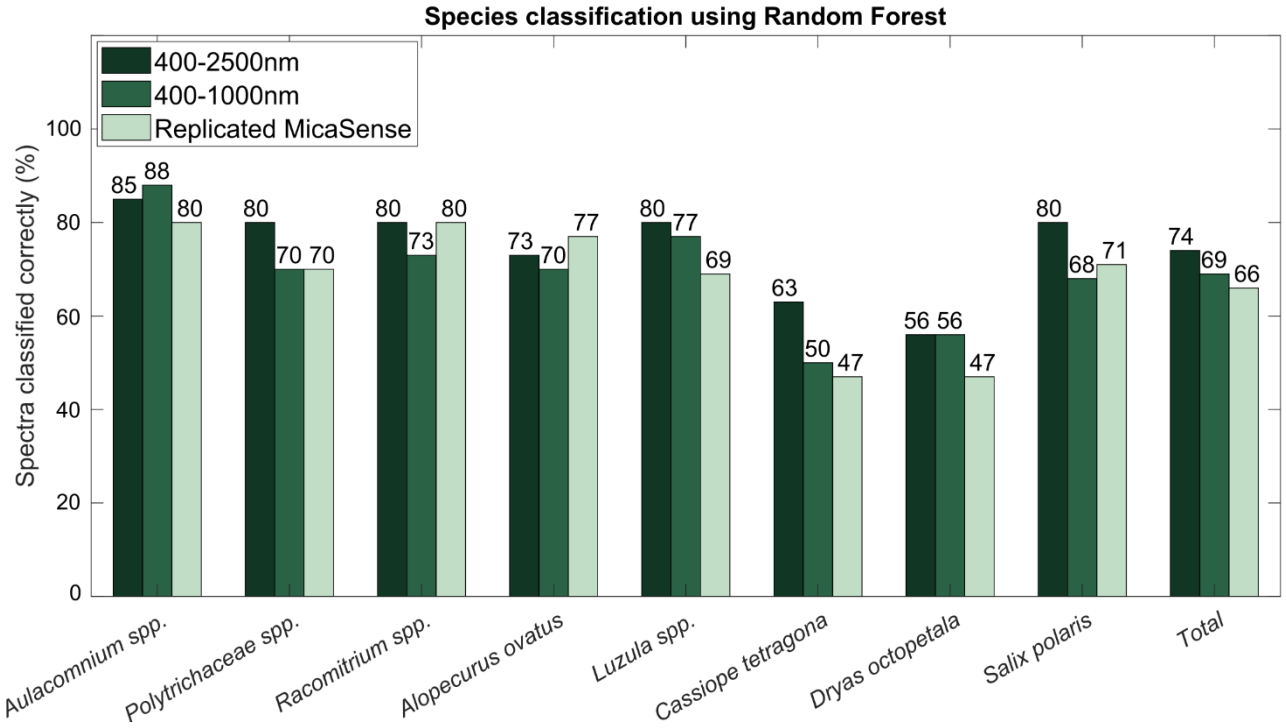


Figure 5.5: Percentage of turf spectra correctly classified using Random Forest classification. Five spectra per turf were used (n = 265). The spectra were classified three times: (1) using the full hyperspectral range from 400 to 2500 nm, (2) just the VIS–NIR region from 400 to 1000 nm and (3) downscaled spectra to match the five bands represented by the MicaSense-MX camera (blue, green, red, red edge and near-infrared).

Field spectroscopy identification of plant traits

Using the full reflectance spectra (400–2500 nm), PLSR results show that plant water content was estimated with the highest accuracy (r^2 = 0.88, RMSE = 12%), followed by N (r^2 = 0.77, RMSE = 28%), P (r^2 = 0.67, RMSE = 64%) and C:N ratios (r^2 = 0.61, RMSE = 37%, Figure **5.6(a)**). Removing spectral information from the SWIR range had the greatest impact on water prediction. Decreased SWIR spectral information had almost no impact on N prediction, and little impact on C:N ratio prediction. P estimation showed little change with reduced SWIR information but was strongly influenced by fewer VIS–NIR spectral bands (**Figures 6(b) and (c)**). The spectral weightings showed that all regions of the spectrum are important to trait prediction when the full spectrum is used, but the red-edge and NIR region are particularly important when SWIR information is lacking (**Figures 6(d)–(f)**).

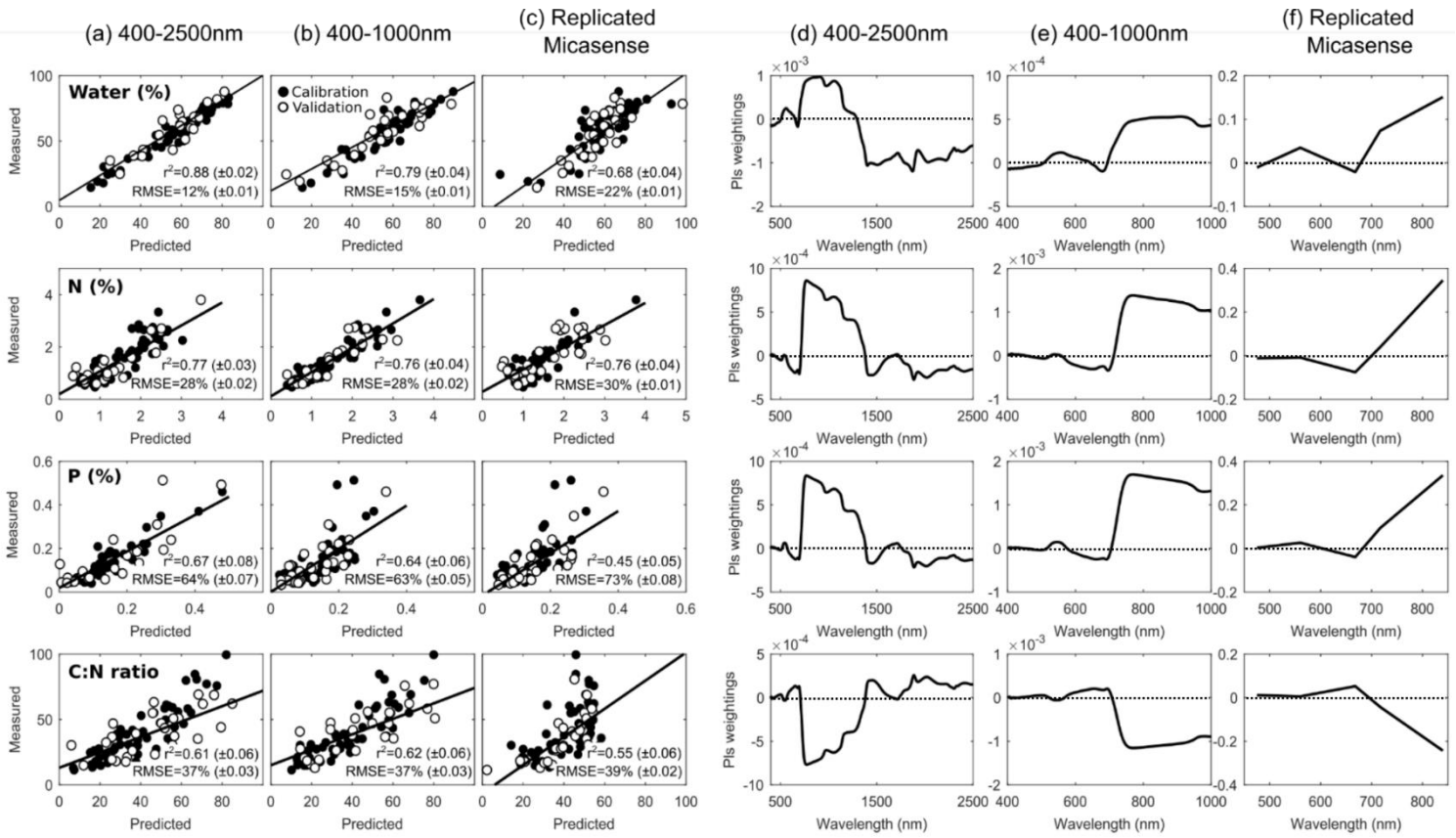


Figure 5.6: (a)–(c) Results of predicted versus measured plant trait values (water (%), nitrogen (%), phosphorus (%) and C:N ratios) using PLSR analysis. (d)–(f) PLSR spectral weightings for each plant trait. Deviation from 0 indicates the parts of the spectrum that contribute most strongly to prediction. Spectra were analysed three times, using: (1) the full hyperspectral range from 400 to 2500 nm, (2) the VIS–NIR region from 400 to 1000 nm and (3) downscaled spectra to simulate the five bands represented by the MicaSense-MX camera (blue, green, red, red edge and near-infrared).

UAV spectral mapping of plant functional types

When using the UAV MicaSense camera, the eight vegetation classes could not be classified with the same accuracy as when using the downscaled hyperspectral data. Instead, vegetation cover was classified into three plant functional types; mosses, graminoids and dwarf shrubs (**Figure 5.7(b)**). Overall classification accuracies were 72% across the validation dataset, much above the statistical random accuracy of 33% expected for three classes. Dwarf shrubs were classified with the highest accuracy (83%), followed by graminoids (61%) and mosses (57%, **Figure 5.7(c)**). Vegetation cover across the Birdcliff site was dominated by dwarf shrubs. Graminoids were shown as present at the confluence of runoff channels and on some of the flatter, wetter areas before the fjord. Mosses showed speckling between the shrubs and were found in dense communities around the bottom of the colluvial fans.

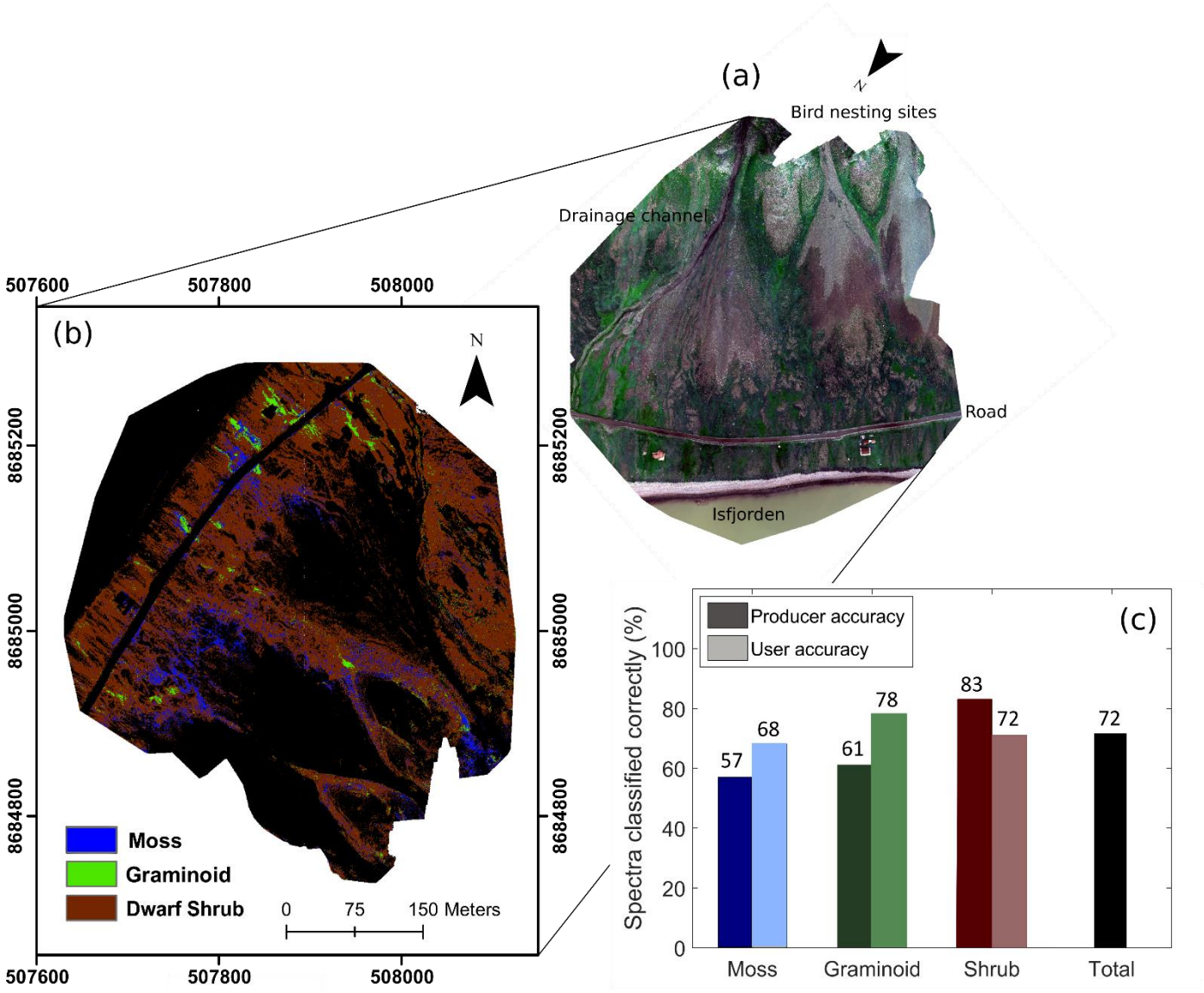


Figure 5.7: (a) Annotated RGB image of the Birdcliff site. The image has been rotated to represent decreasing elevation from top to bottom. (b) Map of moss, graminoid and dwarf shrub distribution across the Birdcliff site. Produced using UAV reflectance spectra and Random Forest Classification. (c) User and producer accuracy of the model applied to the validation dataset.

UAV spectral mapping of plant traits

Using the UAV five-band data, average correlations with water and N were significantly lower ($p < 0.001$) than when using the downscaled hyperspectral data (**Figure 5.8**). Plant water trait correlations decreased from $r^2 = 0.68$ to $r^2 = 0.50$ and N accuracies decreased from $r^2 = 0.76$ to $r^2 = 0.50$. The exception was P, where the average correlations between the UAV spectroscopy and field measurements increased from $r^2 = 0.45$ to $r^2 = 0.54$ compared to the downscaled hyperspectral data. Scaling up the most robust models, there was a clear plant water and nutrient gradient across the site. Plant water and nutrient values were generally highest at the top of the nutrient and elevation gradient and declined downslope. Nutrient concentrations also appeared to follow drainage channels across the site and were tightly correlated to each other.

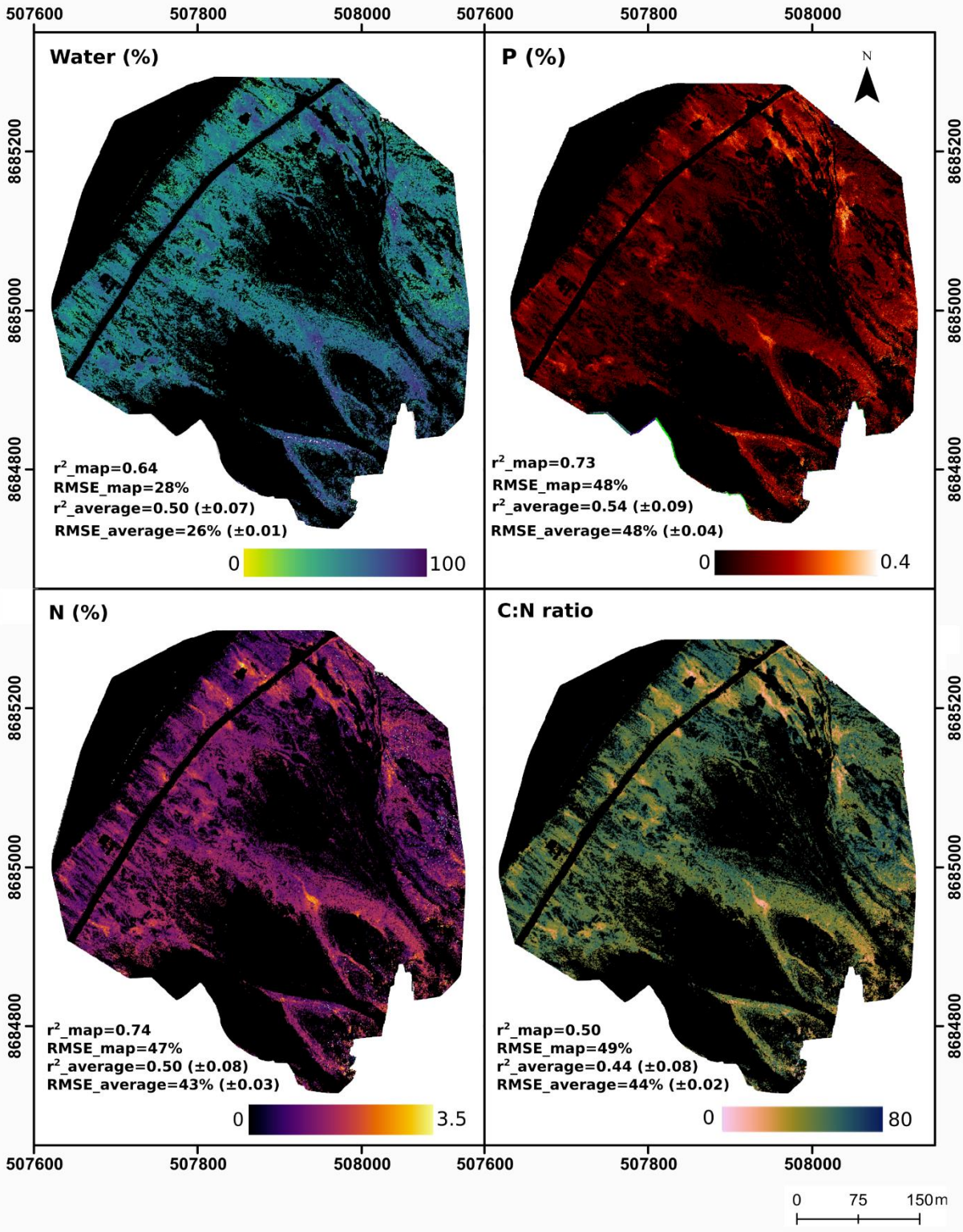


Figure 5.8: Maps of plant water, nitrogen, phosphorus and C:N ratios across the Birdcliff site. The maps were produced using UAV spectra and partial least squares regression PLSR). r^2_map and RMSE_map represent the PLSR model that was chosen to scale up the results of the calibration/validation dataset. r^2_average and RMSE_average are the average of 1000 PLSR runs.

Sentinel-2 spectral mapping of plant water content

Scaling from the UAV to the satellite level, calibrating a Sentinel-specific PLSR model on the UAV plant water map values (**Figure 5.9(a)**) and applying the UAV plant water model directly to the Sentinel imagery (**Figure 5.9(b)**) produced similar patterns but a very different magnitude of result. The maps were correlated with each other at $r = 0.86$, with a y-intercept of 41%. For method A, plant water content values ranged between 30% and 80%, while method B overestimated water content, with values ranging from 80% upwards.

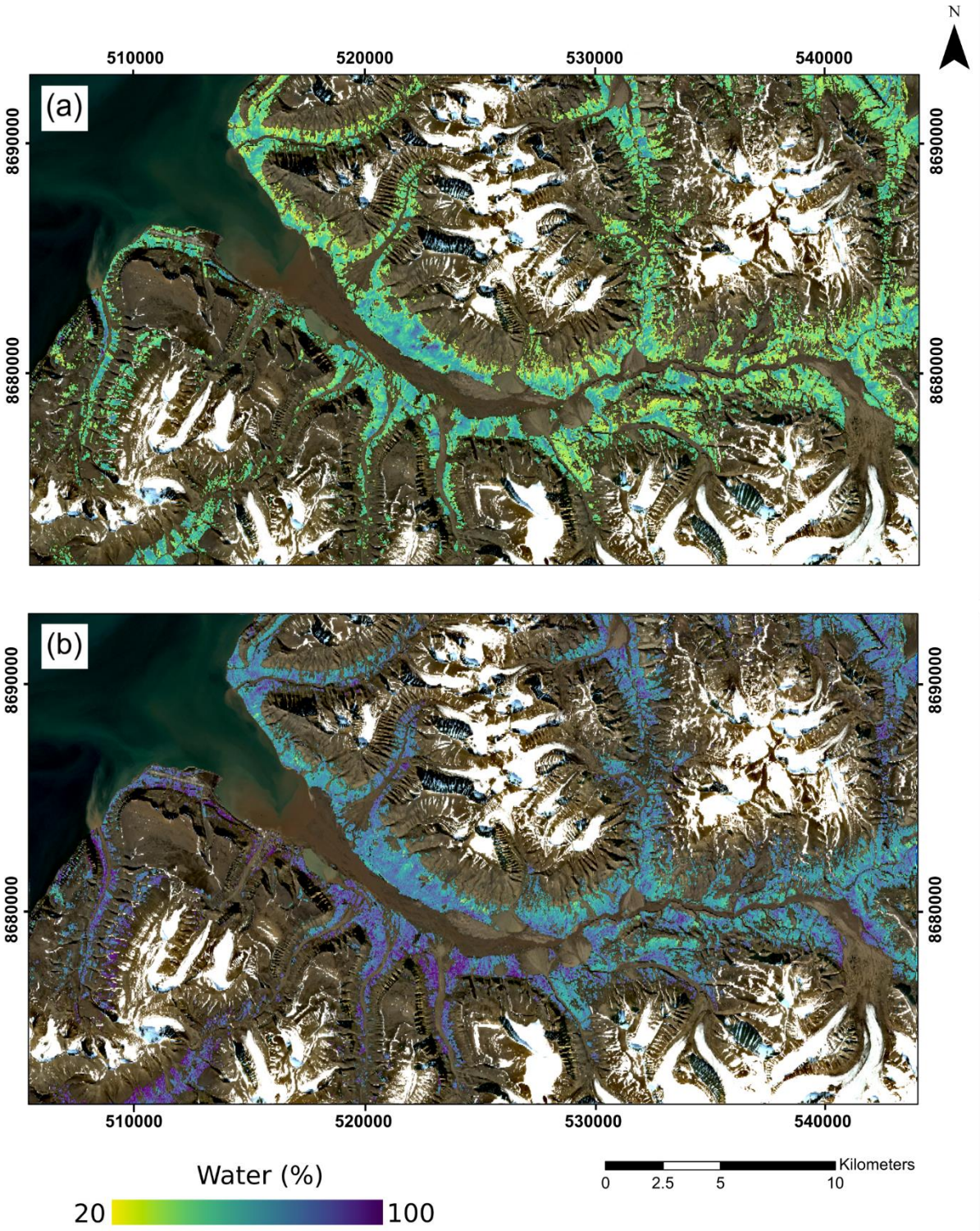


Figure 5.9: Two methods of mapping vegetation water content (%) around Longyearbyen, Svalbard using the UAV-generated plant water map in **Figure 5.8** and Sentinel-2 imagery. Method (a): calibrating a Sentinel-specific PLSR model (ten bands) on the UAV plant water map values and applying it across the imagery. Method (b): applying the UAV plant water PLSR model (five bands) to the closest corresponding Sentinel-2 bands. The cloud-free Sentinel-2 image was taken on 30 July 2018.

5.7. Discussion

Our results show that UAV multispectral data can be used to map fine-scale vegetation cover (moss, graminoids and dwarf shrubs) across the High Arctic and monitor changes in plant water and nutrient traits. Using field spectroscopy data, we were able to distinguish eight common tundra species at 74% accuracy, suggesting good prospects for near-future vegetation mapping at the species level with the increasing commercialisation of UAV–hyperspectral systems and next generation of hyperspectral satellites (CHIME, EnMAP, HISUI, HypXIM, HyspIRI, PRISMA: Nieke and Rast, 2018; Guanter *et al.* 2016; Matsunaga *et al.* 2013; Carrere *et al.* 2013; Lee *et al.* 2015; Pignatti *et al.* 2013). Although only CHIME, EnMAP and PRISMA cover the high latitudes (extending to 84° N, 80° N and 70° N respectively), hyperspectral technology in this area is expanding.

Mosses in particular displayed a unique spectral signature and distinct spectral separability within tundra vegetation. Average moss reflectance across the NIR and SWIR was lower than for graminoids and shrubs, reflecting their lack of vascular system and high water content. All three families/genera of moss, *Aulacomnium, Polytrichaceae* and *Racomitrium*, displayed high prediction accuracies in the Random Forest analysis (85, 80% and 80% respectively), which is promising for the automatic detection of moss taxa in future hyperspectral studies. In northern environments, mosses can dominate aboveground primary productivity and have been shown to display a wide range of traits and life history strategies (Longton, 1988; Hedderson and Longton, 1996; Turetsky *et al*. 2010; Elumeeva *et al.* 2011). Yet, due to the difficulty of identifying species, mosses are often ignored by ecologists in favour of vascular plants, or conglomerated into groups that may or may not be meaningful (Turetsky *et al*. 2012). Thus, the ability to distinguish moss taxa at broad scales would facilitate research focusing on this significant but understudied plant group.

Graminoids showed similar spectral separability with a 73% and 80% prediction accuracy for *Alopecurus ovatus* and *Luzula*, respectively. *Salix polaris* also demonstrated a high prediction accuracy of 80%, whereas some spectral confusion was created between the two evergreen dwarf shrubs, *Cassiope* and *Dryas*, likely due to *Cassiope*'s

woody vertical structure that created high variation in its spectral signature (**Figure 5.3**). Developing spectroscopic techniques to improve the identification of dwarf shrub species would be useful at broad scales as, along with snow depth, the litter quality of *Cassiope, Dryas* and *Salix* has been linked to differences in soil organic matter and respiration rates (Elberling, 2007). *Cassiope, Dryas* and *Salix* also display contrasting evergreen and deciduous strategies, which govern different relationships between soil nutrient availability, photosynthesis and growth rates (Baddeley *et al.* 1994). Overall, these results are highly encouraging with respect to the spectral separability of tundra vegetation and application of fine-scale remote sensing in the Arctic.

At the trait level, we found that four plant traits (water, N, P and C:N values) could be estimated with varying accuracy using field spectroscopy. Plant water and N concentrations were strongly correlated with the full range spectra and showed good agreement with previous PLSR studies at the leaf and canopy scale (e.g. Asner *et al*. 2015; Asner and Martin, 2008; Ullah *et al.* 2014). The estimation of P was slightly lower than previous studies (e.g. Mutanga and Kumar, 2007; Baddeley *et al.* 1994; Asner *et al.* 2014; Asner *et al.* 2015,), although the PLSR models still captured a substantial amount of variation across the data. Unlike water and N, P has no direct expression across the 400–2500 nm spectrum. Predictions result from correlations with other traits, especially nitrogen, attributed to the stoichiometric link between N and P—a link which has been used to map leaf P across large areas of the tropics and identify crop-level P deficiencies (Porder *et al.* 2005; Mahajan *et al.* 2014; Asner *et al.* 2016). In contrast to forest or agricultural studies however, N:P ratios may differ more widely among our samples, through different plant groups having fundamentally different N:P ratios (Mahajan *et al.* 2017) and high nutrient deposition from bird colonies leading to uneven N:P soil ratios.

When downscaling the hyperspectral data, estimations of plant water content showed a corresponding decrease in accuracy with the number of reduced spectral bands. The nutrient traits were less sensitive to reduced SWIR information, reflecting that water is expressed in a wide range of spectral bands across the VNIR–SWIR spectrum, whereas N is a major component of chlorophyll with strong expression in blue and green VIS

bandwidths (430–660 nm (Curran *et al.* 1989; Kokaly, 2001). P displayed similar spectral weightings to N although, when downscaled to the five MicaSense bands, P values were highly sensitive to a reduction in VNIR information (as similarly found by (Milton *et al.* 1991; Mutanga and Kumar, 2007). In contrast, N showed no significant reduction in accuracy across the downscaled analyses, further suggesting that, in Arctic environments, estimations of P arise from collinearities with traits other than N. We have confidence that the models are detecting trait variation, rather than species-level differences, given the observed spectral weightings and high intraspecific trait variation within Arctic plant functional groups (see **Table S5.1** and Thomas *et al.* 2019).

Using the multispectral UAV data, we were able to map moss, graminoids and dwarf shrubs at 72% accuracy. Forbs were not included in the model, as the vegetation cover of any individual forb species was not extensive enough to sample sufficiently. In future studies, it would be useful to test whether forbs can be spectrally distinguished from graminoids (as found by Langford *et al*. 2016) using multispectral satellite imagery, LiDAR and phenology), or whether their similar functionality render forbs and graminoids too related to be differentiated by spectra alone.

Across the Birdcliff site, modelled vegetation distribution was dominated by shrub cover. High nutrient input from nesting colonies generally leads to the formation of moss tundra below bird cliffs on Svalbard (Hofman, 1986; Vanderpuye *et al.* 2002), but along warmer coastlines, such as Bjørndalen, vascular plants have been found to proliferate (Elven *et al.* 1990). Although moss only constituted 10% of the vegetation cover across the site, the ability to separate moss-dominated pixels from shrubs and graminoids is a key strength of high-resolution mapping over satellite-scale analyses. Regular monitoring of moss communities can act as an early warning system for climatic shifts and associated biotic and abiotic effects in polar environments (Callaghan *et al.* 1999; Robinson *et al.* 2005; Royles *et al.* 2013). However, it is important to remember that this study represented a snapshot in time, where moisture conditions, illumination conditions and phenological phase were unique. Time-series data is required to investigate whether moss and other vegetation types can be distinguished with the same accuracy at different phenological stages and under varying abiotic conditions (Van Leeuwen and Huete, 1996; May *et al.* 2018).

Using the UAV imagery, we were unable to distinguish vegetation at the species level, despite the downscaled hyperspectral data displaying species level accuracy of 64%. We attribute this loss in accuracy to two main reasons. Firstly, UAV imagery suffers from quality issues to a greater extent than field-measured spectra. These include bidirectional reflectance distribution function effects, variability in illumination conditions during a flight, poor radiometric calibration and spectral mixing, even at high resolutions (Assmann *et al.* 2018; Stow *et al.* 2019). These effects are even more pronounced in the Arctic due to low sun angle, variable cloud cover and heterogeneous tundra. While the UAV imagery shows a correlation with the field-collected spectra (**Figure S5.2**), there is substantial variation in the agreement, with the UAV spectra likely exhibiting more statistical error. High variation in the UAV spectra may explain why the highest accuracies in our analyses were achieved by averaging spectra from a square of nine pixels (18 × 18 cm) rather than extracting single pixels, due to the increased signal-to-noise ratio of the averaged spectra.

Secondly, accurate geolocation of the ground-truthed vegetation points can be challenging. Across the Arctic, species distribution is extremely heterogeneous and ground-truthing points may only define species presence within a radius of <20 cm. Although RMS errors for the orthomosaics were low (<0.14 m), geolocation error could still have resulted in an offset between the ground-truthed vegetation points and the imagery, leading to some inaccuracies in the calibration and validation points. Thus, in environments with high spatial heterogeneity, extremely precise georeferencing equipment is required to accurately calibrate and validate the information collected from UAVs. We recommend covering vegetation around ground-truthing points to create an exclusion zone, or similar methods to reduce the spectral effect of nearby species. In addition, when dealing with mixed plant communities, novel spectral methods are required to capture the co-occurrence of plant types. Rather than discrete vegetation categories, a more flexible model may be required, that removes individual classifications and instead places a vegetation pixel on a continuum of values between moss, graminoid or shrub.

At the landscape level, plant trait distributions displayed a steep water and fertility gradient from the bird cliff colony to the sea, mediated by geomorphology. Plant water

content was highest directly underneath the bird colonies and at the confluence of drainage channels. N, P and C:N values indicated well-fertilised vegetation under nesting sites and reflected the transport of nutrients downslope through a network of drainage channels. For example, plant water and nutrient levels were extremely high in the depositional area below the snowmelt channel (**Figure 5.7**), demonstrating the role that run-off processes play in distributing marine-derived organic matter across the otherwise nutrient-limited tundra. While graminoids constituted just 4% of the vegetation cover across the mapped area, they were mostly found in moist, fertile areas, reflecting their fast growth rates and ability to take swift advantage of available nutrients (Wookey *et al.* 2009). The dominance of gramanids in these areas suggests that increased mineralization rates under a warming climate may preferentially benefit graminoid species (as found by Theodose and Bowman, 1997; Tomassen, 2004; Van Der Wal and Brooker, 2004).

Adding to our confidence that the models were predicting traits explicitly, rather than through correlations with plant type, vegetation cover or productivity, multispectral methods of plant trait retrieval showed a clear independence from NDVI-based predictions (**Figure S5.3**). With the exception of C:N ratios, neither measured nor predicted trait values were significantly correlated with NDVI, demonstrating the value of a multispectral mapping approach and the lack of physical basis for using NDVI to predict plant traits in the Arctic.

Scaling from the UAV to satellite level, we found different methods of upscaling had a large effect on the magnitude of results, but little influence on the spatial pattern of trait values. Although we do not present either map in **Figure 5.9** as accurate, or have the data to validate either map, we found that method A (calibrating Sentinel imagery on UAV-derived plant water values) produced more varying and realistic trait values than method B (applying a UAV plant water model to Sentinel imagery). The higher variation in trait values using method A is probably due to the greater number of spectral bands used in the analysis. In method A, all the relevant Sentinel bands were included as predictors (ten in total), whereas in method B, only the five Sentinel bands that best matched the spectral bands of the MicaSense camera could be used by the UAV plant water model. The overestimation of trait values in method B demonstrates the

difference in reflectance captured by UAVs and satellites (see **Figure S5.4**). Cross-sensor differences arise from atmospheric effects (e.g. haze), differences in viewing and illumination angles, sensor specifications (e.g. bandwidths), and non-linear spectral mixing at different spatial resolutions (Bucchorn *et al.* 2016; Stow *et al.* 2019; Beamish *et al.* 2020; Myers-Smith *et al.* 2020), meaning predictive models generated from one type of imagery cannot be directly applied to another with the same accuracy.

To upscale from UAV to satellite imagery therefore, we recommend creating satellite-specific models calibrated on UAV maps (e.g. method A). Similar methods have been used for a variety of ecological-mapping tasks, such as tree species classification, percentage vegetation cover, agricultural assessments, habitat mapping and fire severity quantification (Immitzer *et al.* 2012; Asner *et al.* 2015; Fraser *et al.* 2017; Pla *et al*. 2017; Leitão *et al.* 2018; Pla *et al*. 2019; Kattenborn *et al.* 2019; Riihimäki *et al.* 2019). To the best of our knowledge however, this study is the first time that upscaling from UAV to satellite imagery has been used to quantify plant traits. With greater UAV coverage across the landscape and more accurate UAV-models ($r^2 > 0.9$), using UAV information to calibrate satellite-based models may provide a valuable method for expanding spatially limited plot studies across the Arctic.

5.8. Conclusion

Using field spectroscopy, we showed that eight common Arctic plants can be distinguished at the species level at 74% accuracy. These results are relevant for UAV-hyperspectral systems, the first Arctic passenger aircraft equipped with hyperspectral instruments (Sivertson *et al.* 2019) and upcoming hyperspectral satellites, which should be able to use similar methods explored here to advance these results and expand them over a wider geographical area.

Using UAV multispectral information, we have shown that High Arctic moss-, graminoid- and dwarf shrub-dominated vegetation can be mapped at high resolution (<10 cm) with 72% accuracy, alongside corresponding water, nitrogen, phosphorus and C:N values. We present two methods of upscaling these results using Sentinel-2 imagery. The majority of Arctic tundra regions remain under-investigated and difficult to access for scientific research. We must continue to develop methods to expand spatially-limited plot studies, which inherently cannot capture landscape-level vegetation patterns or functional trait variation, in order to monitor environmental change across the Arctic and understand the mechanisms driving large-scale climate-vegetation feedbacks.

5.9. Acknowledgments

We thank Plant Functional Trait Course 4 for the funding and logistics that enabled the data collection. We thank the course instructors and course participants for their help and collaboration during the data collection and processing. We also thank Christine Schirmer and the students at the University of Arizona for their assistance with the chemical analyses.

5.10. Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://osf.io/smbqh/.

5.11. Funding

The data collection was funded by a Norwegian Research Council INTPART Grant (Project Number: 274831), two SIU-foundation projects (UTF-2013/10074 and HNP-2015/10037) and a Research Council of Norway Arctic Field Grant (Project Number: 282611, RiS: 10935). E R T is funded by NERC DTP award (NE/L002612/1). Y M is supported by the Jackson Foundation. M M-F was supported by a NERC IRF (NE/L011859/1).

5.12 References

Arndt, K.A., Santos, M.J., Ustin, S., Davidson, S.J., Stow, D., Oechel, W.C., Tran, T.T.P., Graybill, B. and Zona, D. (2019) Arctic greening associated with lengthening growing seasons in Northern Alaska. *Environmental Research Letters,* 14 (12), pp. 125018.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J. and Hughes, R.F. (2008) Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sensing of Environment,* 112 (5), pp. 1942-1955.

Asner, G.P., Martin, R.E. and Suhaili, A.B. (2012) Sources of canopy chemical and spectral diversity in lowland Bornean forest. *Ecosystems,* 15 (3), pp. 504-517.

Asner, G.P., Martin, R.E., Carranza-Jiménez, L., Sinca, F., Tupayachi, R., Anderson, C.B. and Martinez, P. (2014) Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to Amazon region. *New Phytologist,* 204 (1), pp. 127-139.

Asner, G.P., Martin, R.E., Anderson, C.B. and Knapp, D.E. (2015) Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sensing of Environment,* 158, pp. 15-27.

Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E. and Vaughn, N. (2016) Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proceedings of the National Academy of Sciences,* 113 (28), pp. E4043-E4051.

Asner, G.P. and Martin, R.E. (2008) Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment,* 112 (10), pp. 3958-3970.

Asner, G.P. and Martin, R.E. (2009) Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment,* 7 (5), pp. 269-276.

Assmann, J.J., Kerby, J.T., Cunliffe, A.M. and Myers-Smith, I. (2018) Vegetation monitoring using multispectral sensors—Best practices and lessons learned from high latitudes. *Journal of Unmanned Vehicle Systems,* 7 (1), pp. 54-75.

Assmann, J.J., Myers-Smith, I., Kerby, J.T., Cunliffe, A.M. and Daskalova, G.N. (2020) Drone data reveal heterogeneity in tundra greenness and phenology not captured by satellites. *Environmental Research Letters,* 15 (12), pp. 125002.

Axelsson, C., Skidmore, A.K., Schlerf, M., Fauzi, A. and Verhoef, W. (2013) Hyperspectral analysis of mangrove foliar chemistry using PLSR and support vector regression. *International Journal of Remote Sensing,* 34 (5), pp. 1724-1743

Baddeley, J.A., Woodin, S.J. and Alexander, I.J. (1994) Effects of increased nitrogen and phosphorus availability on the photosynthesis and nutrient relations of three arctic dwarf shrubs from Svalbard. *Functional Ecology,* , pp. 676-685.

Baena, S., Moat, J., Whaley, O. and Boyd, D.S. (2017) Identifying species from the air: UAVs and the very high resolution challenge for plant conservation. *PloS one,* 12 (11).

Beamish, A.L., Coops, N., Chabrillat, S. and Heim, B. (2017) A phenological approach to spectral differentiation of low-arctic tundra vegetation communities, north slope, Alaska. *Remote Sensing,* 9 (11), pp. 1200.

Beamish, A., Raynolds, M.K., Epstein, H., Frost, G.V., Macander, M.J., Bergstedt, H., Bartsch, A., Kruse, S., Miles, V. and Tanis, C.M. (2020) Recent trends and remaining

challenges for optical remote sensing of Arctic tundra vegetation: A review and outlook. *Remote Sensing of Environment,* 246, pp. 111872.
Bhatt, U.S., Walker, D.A., Raynolds, M.K., Bieniek, P.A., Epstein, H.E., Comiso, J.C., Pinzon, J.E., Tucker, C.J. and Polyakov, I.V. (2013) Recent declines in warming and vegetation greening trends over pan-Arctic tundra. *Remote Sensing,* 5 (9), pp. 4229-4254.
Bjerke, J.W., Treharne, R., Vikhamar-Schuler, D., Karlsen, S.R., Ravolainen, V., Bokhorst, S., Phoenix, G.K., Bochenek, Z. and Tømmervik, H. (2017) Understanding the drivers of extensive plant damage in boreal and Arctic ecosystems: Insights from field surveys in the aftermath of damage. *Science of the Total Environment,* 599, pp. 1965-1976.
Bjorkman, A.D., Myers-Smith, I., Elmendorf, S.C., Normand, S., Rüger, N., Beck, P.S.A., Blach-Overgaard, A., Blok, D., Cornelissen, J.H. and Forbes, B.C. (2018) Plant functional trait change across a warming tundra biome. *Nature,* 562 (7725), pp. 57-62.
Blok, D., Schaepman-Strub, G., Bartholomeus, H., Heijmans, M.M.P.D., Maximov, T.C. and Berendse, F. (2011) The response of Arctic vegetation to the summer climate: relation between shrub cover, NDVI, surface albedo and temperature. *Environmental Research Letters,* 6 (3), pp. 035502.
Bokhorst, S., Tømmervik, H., Callaghan, T.V., Phoenix, G.K. and Bjerke, J.W. (2012) Vegetation recovery following extreme winter warming events in the sub-Arctic estimated using NDVI from remote sensing and handheld passive proximal sensors. *Environmental and experimental botany,* 81, pp. 18-25.
Bonfils, C., Phillips, T.J., Lawrence, D.M., Cameron-Smith, P., Riley, W.J. and Subin, Z.M. (2012) On the influence of shrub height and expansion on northern high latitude climate. *Environmental Research Letters,* 7 (1), pp. 015503.
Box, J.E., Colgan, W.T., Christensen, T.R., Schmidt, N.M., Lund, M., Parmentier, F.W., Brown, R., Bhatt, U.S., Euskirchen, E.S. and Romanovsky, V.E. (2019) Key indicators of Arctic climate change: 1971–2017. *Environmental Research Letters,* 14 (4), pp. 045010.
Bratsch, S.N., Epstein, H.E., Buchhorn, M. and Walker, D.A. (2016) Differentiating among four Arctic tundra plant communities at Ivotuk, Alaska using field spectroscopy. *Remote Sensing,* 8 (1), pp. 51.
Bratsch, S., Epstein, H., Buchhorn, M., Walker, D. and Landes, H. (2017) Relationships between hyperspectral data and components of vegetation biomass in Low Arctic tundra communities at Ivotuk, Alaska. *Environmental Research Letters,* 12 (2), pp. 025003.
Breiman, L. (2001) Random forests. *Machine Learning,* 45 (1), pp. 5-32.
Buchhorn, M., Raynolds, M.K. and Walker, D.A. (2016) Influence of BRDF on NDVI and biomass estimations of Alaska Arctic tundra. *Environmental Research Letters,* 11 (12), pp. 125002.
Buchhorn, M., Walker, D.A., Heim, B., Raynolds, M.K., Epstein, H.E. and Schwieder, M. (2013) Ground-based hyperspectral characterization of Alaska tundra vegetation along environmental gradients. *Remote Sensing,* 5 (8), pp. 3971-4005.
Callaghan, T.V., Press, M.C., Lee, J.A., Robinson, D.L. and Anderson, C.W. (1999) Spatial and temporal variability in the responses of Arctic terrestrial ecosystems to environmental change. *Polar Research,* 18 (2), pp. 191-197.
Capolupo, A., Kooistra, L., Berendonk, C., Boccia, L. and Suomalainen, J. (2015) Estimating plant traits of grasslands from UAV-acquired hyperspectral images: a comparison of statistical approaches. *ISPRS International Journal of Geo-Information,* 4 (4), pp. 2792-2820.

Carrere, V., Briottet, X., Jacquemoud, S., Marion, R., Bourguignon, A., Chami, M., Dumont, M., Minghelli-Roman, A., Weber, C. and Lefevre-Fonollosa, M. (2013) HYPXIM: a second generation high spatial resolution hyperspectral satellite for dual applications. IEEE, pp. 1.

Constantin, D., Rehak, M., Akhtman, Y. and Liebisch, F. (2015) Detection of crop properties by means of hyperspectral remote sensing from a micro UAV. *Bornimer Agrartechnische Berichte,* 88, pp. 129-137.

Cornwell, W.K., Cornelissen, J.H.C., Amatangelo, K., Dorrepaal, E., Eviner, V.T., Godoy, O., Hobbie, S.E., Hoorens, B., Kurokawa, H. and Pérez-Harguindeguy, N. (2008) Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecology Letters,* 11 (10), pp. 1065-1071.

Curran P J (1989) Remote sensing of foliar chemistry *Remote Sens. Environ.* 30 271–8

Dash, J.P., Pearse, G.D. and Watt, M.S. (2018) UAV multispectral imagery can complement satellite data for monitoring forest health. *Remote Sensing,* 10 (8), pp. 1216.

Davidson, S.J., Santos, M.J., Sloan, V.L., Watts, J.D., Phoenix, G.K., Oechel, W.C. and Zona, D. (2016a) Mapping Arctic tundra vegetation communities using field spectroscopy and multispectral satellite data in North Alaska, USA. *Remote Sensing,* 8 (12), pp. 978.

Davidson, S.J., Sloan, V.L., Phoenix, G.K., Wagner, R., Fisher, J.P., Oechel, W.C. and Zona, D. (2016b) Vegetation type dominates the spatial variability in CH 4 emissions across multiple arctic tundra landscapes. *Ecosystems,* 19 (6), pp. 1116-1132.

Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C. and Prentice, I.C. (2016) The global spectrum of plant form and function. *Nature,* 529 (7585), pp. 167-171.

Durgante, F.M., Higuchi, N., Almeida, A. and Vicentini, A. (2013) Species spectral signature: discriminating closely related plant species in the Amazon with near-infrared leaf-spectroscopy. *Forest Ecology and Management,* 291, pp. 240-248.

Elberling, B. (2007) Annual soil CO2 effluxes in the High Arctic: the role of snow thickness and vegetation type. *Soil Biology and Biochemistry,* 39 (2), pp. 646-654.

Elumeeva, T.G., Soudzilovskaia, N.A., During, H.J. and Cornelissen, J.H.C. (2011) The importance of colony structure versus shoot morphology for the water balance of 22 subarctic bryophyte species. *Journal of Vegetation Science,* 22 (1), pp. 152-164.

Elven R, Eriksen M B, Elvebakk A, Johansen B and Engelskjøn T (1990) Gipsdalen, central svalbard; flora, vegetation and botanical values *Environmental Atlas, Gipsdalen, Svalbard. II. Nor. Polarinst. Rapp. Ser.* vol 61 27–66

Fauchald, P., Park, T., Tømmervik, H., Myneni, R. and Hausner, V.H. (2017) Arctic greening from warming promotes declines in caribou populations. *Science Advances,* 3 (4), pp. e1601365.

Fraser, R.H., Olthof, I., Lantz, T.C. and Schmitt, C. (2016) UAV photogrammetry for mapping vegetation in the low-Arctic. *Arctic Science,* 2 (3), pp. 79-102.

Fraser, R.H., Van der Sluijs, J. and Hall, R.J. (2017) Calibrating satellite-based indices of burn severity from UAV-derived metrics of a burned boreal forest in NWT, Canada. *Remote Sensing,* 9 (3), pp. 279.

Frost, G.V., Epstein, H.E., Walker, D.A., Matyshak, G. and Ermokhina, K. (2018) Seasonal and long-term changes to active-layer temperatures after tall shrubland expansion and succession in Arctic tundra. *Ecosystems,* 21 (3), pp. 507-520.

Ghimire, B., Rogan, J. and Miller, J. (2010) Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sensing Letters,* 1 (1), pp. 45-54.
Guanter, L., Kaufmann, H., Foerster, S., Brosinsky, A., Wulf, H., Bochow, M., Boesche, N., Brell, M., Buddenbaum, H. and Chabrillat, S. (2016) *EnMAP Science Plan—Environmental Mapping and Analysis Program,* .
Guay, K.C., Beck, P.S.A., Berner, L.T., Goetz, S.J., Baccini, A. and Buermann, W. (2014) Vegetation productivity patterns at high northern latitudes: A multi-sensor satellite data assessment. *Global Change Biology,* 20 (10), pp. 3147-3158.
Hedderson, T.A. and Longton, R.E. (1996) Life history variation in mosses: water relations, size and phylogeny. *Oikos,* , pp. 31-43.
Henry, G. and Molau, U. (1997) Tundra plants and climate change: the International Tundra Experiment (ITEX). *Global Change Biology,* 3, pp. 1-9.
Higuera, P.E., Brubaker, L.B., Anderson, P.M., Brown, T.A., Kennedy, A.T. and Hu, F.S. (2008) Frequent fires in ancient shrub tundra: implications of paleorecords for arctic environmental change. *PloS one,* 3 (3).
Hijmans, R.J., Van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J.A., Lamigueiro, O.P., Bevan, A., Racine, E.B. and Shortridge, A. (2015) Package 'raster'. *R package,* .
Hofmann, W. (1968) Geobotanische Untersuchungen in Sudost-Spitzbergen 1960.
Hudson, J.M.G. and Henry, G.H.R. (2009) Increased plant biomass in a High Arctic heath community from 1981 to 2008. *Ecology,* 90 (10), pp. 2657-2663.
Huemmrich, K.F., Gamon, J.A., Tweedie, C.E., Campbell, P.K.E., Landis, D.R. and Middleton, E.M. (2013) Arctic tundra vegetation functional types based on photosynthetic physiology and optical properties. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 6 (2), pp. 265-275.
Immitzer, M., Atzberger, C. and Koukal, T. (2012) Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote sensing,* 4 (9), pp. 2661-2693.
Ims, R.A., Henden, J., Strømeng, M.A., Thingnes, A.V., Garmo, M.J. and Jepsen, J.U. (2019) Arctic greening and bird nest predation risk across tundra ecotones. *Nature Climate Change,* 9 (8), pp. 607-610.
Jones, A., Stolbovay, V., Tarnocai, C., Broll, G., Spaargaren, O. and Montanarella, L. (2010) *Soil atlas of the Northern Circumpolar Region.* European Commission.
Joshua L A, Klein E S, Oberbauer S F and Welker J M (2016) Coupled long-term summer warming and deeper snow alters species composition and stimulates gross primary productivity in tussock tundra *Oecologia* 181 287–97
Juszak, I., Erb, A.M., Maximov, T.C. and Schaepman-Strub, G. (2014) Arctic shrub effects on NDVI, summer albedo and soil shading. *Remote Sensing of Environment,* 153, pp. 79-89.
Kattenborn, T., Lopatin, J., Förster, M., Braun, A.C. and Fassnacht, F.E. (2019) UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data. *Remote Sensing of Environment,* 227, pp. 61-73.
Kennedy, B.E., King, D.J. and Duffe, J. (2020) Comparison of Empirical and Physical Modelling for Estimation of Biochemical and Biophysical Vegetation Properties: Field Scale Analysis across an Arctic Bioclimatic Gradient. *Remote Sensing,* 12 (18), pp. 3073.

Khorsand Rosa, R., Oberbauer, S.F., Starr, G., Parker La Puma, I., Pop, E., Ahlquist, L. and Baldwin, T. (2015) Plant phenological responses to a long-term experimental extension of growing season and soil warming in the tussock tundra of Alaska. *Global Change Biology,* 21 (12), pp. 4520-4532.

Kokaly R F (2001) Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration *Remote Sens. Environ.* 75 153–61

Kotsiantis, S. and Pintelas, P. (2004) Combining bagging and boosting. *International Journal of Computational Intelligence,* 1 (4), pp. 324-333.

Koven, C.D., Ringeval, B., Friedlingstein, P., Ciais, P., Cadule, P., Khvorostyanov, D., Krinner, G. and Tarnocai, C. (2011) Permafrost carbon-climate feedbacks accelerate global warming. *Proceedings of the National Academy of Sciences,* 108 (36), pp. 14769-14774.

Langford, Z., Kumar, J., Hoffman, F.M., Norby, R.J., Wullschleger, S.D., Sloan, V.L. and Iversen, C.M. (2016) Mapping Arctic plant functional type distributions in the Barrow Environmental Observatory using WorldView-2 and LiDAR datasets. *Remote Sensing,* 8 (9), pp. 733.

Lawrimore J H, Ray R, Applequist S, Korzeniewski B and Menne M J (2016) Global Summary of the Month (GSOM), version 1.0. NOAA National Centers for Environmental Information (Accessed 26 October 2020)

Lee, H., Wang, J. and Leblon, B. (2020) Intra-Field Canopy Nitrogen Retrieval from Unmanned Aerial Vehicle Imagery for Wheat and Corn Fields. *Canadian Journal of Remote Sensing,* , pp. 1-19.

LeitÃ£o, P.J., Schwieder, M., PÃ¶tzschner, F., Pinto, J.R.R., Teixeira, A.M.C., Pedroni, F., Sanchez, M., Rogass, C., van der Linden, S. and Bustamante, M.M.C. (2018) From sample to pixel: multiâ€scale remote sensing data for upscaling aboveground carbon data in heterogeneous landscapes. *Ecosphere,* 9 (8), pp. e02298.

Liu, N., Budkewitsch, P. and Treitz, P. (2017) Examining spectral reflectance features related to Arctic percent vegetation cover: Implications for hyperspectral remote sensing of Arctic tundra. *Remote Sensing of Environment,* 192, pp. 58-72.

Longton, R.E. (1988) Adaptations and strategies of polar bryophytes. *Botanical Journal of the Linnean Society,* 98 (3), pp. 253-268.

Lopatin, J., Fassnacht, F.E., Kattenborn, T. and Schmidtlein, S. (2017) Mapping plant species in mixed grassland communities using close range imaging spectroscopy. *Remote Sensing of Environment,* 201, pp. 12-23.

Loranty, M.M. and Goetz, S.J. (2012) Shrub expansion and climate feedbacks in Arctic tundra. *Environmental Research Letters,* 7 (1), pp. 011005.

Loranty, M.M., Goetz, S.J. and Beck, P.S.A. (2011) Tundra vegetation effects on pan-Arctic albedo. *Environmental Research Letters,* 6 (2), pp. 024014.

Lu, B. and He, Y. (2017) Species classification using Unmanned Aerial Vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland. *ISPRS Journal of Photogrammetry and Remote Sensing,* 128, pp. 73-85.

Lucieer, A., Turner, D., King, D.H. and Robinson, S.A. (2014) Using an Unmanned Aerial Vehicle (UAV) to capture micro-topography of Antarctic moss beds. *International journal of applied earth observation and geoinformation,* 27, pp. 53-62.

Mekonnen Z A, Riley W J and Grant R F (2018) Accelerated nutrient cycling and increased light competition will lead to 21st century shrub expansion in North American Arctic Tundra *J. Geophys. Res. Biogeosci.* 123 1683–701

Mahajan, G.R., Pandey, R.N., Sahoo, R.N., Gupta, V.K., Datta, S.C. and Kumar, D. (2017) Monitoring nitrogen, phosphorus and sulphur in hybrid rice (Oryza sativa L.) using hyperspectral remote sensing. *Precision agriculture,* 18 (5), pp. 736-761.

Mahajan, G.R., Sahoo, R.N., Pandey, R.N., Gupta, V.K. and Kumar, D. (2014) Using hyperspectral remote sensing techniques to monitor nitrogen, phosphorus, sulphur and potassium in wheat (Triticum aestivum L.). *Precision agriculture,* 15 (5), pp. 499-522.

Matsunaga, T., Iwasaki, A., Tsuchida, S., Tanii, J., Kashimura, O., Nakamura, R., Yamamoto, H., Tachikawa, T. and Rokugawa, S. (2013) Current status of hyperspectral imager suite (HISUI). IEEE, pp. 3510.

May, J.L., Parker, T., Unger, S. and Oberbauer, S.F. (2018) Short term changes in moisture content drive strong changes in Normalized Difference Vegetation Index and gross primary productivity in four Arctic moss communities. *Remote Sensing of Environment,* 212, pp. 114-120.

Metcalfe, D.B., Hermans, T.D.G., Ahlstrand, J., Becker, M., Berggren, M., Björk, R.G., Björkman, M.P., Blok, D., Chaudhary, N. and Chisholm, C. (2018) Patchy field sampling biases understanding of climate change impacts across the Arctic. *Nature ecology & evolution,* 2 (9), pp. 1443-1448.

Milton, N.M., Eiswerth, B.A. and Ager, C.M. (1991) Effect of phosphorus deficiency on spectral reflectance and morphology of soybean plants. *Remote Sensing of Environment,* 36 (2), pp. 121-127.

Mora, C., Vieira, G., Pina, P., Lousada, M. and Christiansen, H.H. (2015) Land cover classification using high-resolution aerial photography in adventdalen, svalbard. *Geografiska Annaler: Series A, Physical Geography,* 97 (3), pp. 473-488.

Mutanga, O. and Kumar, L. (2007) Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. *International Journal of Remote Sensing,* 28 (21), pp. 4897-4911.

Myers-Smith, I.H., Thomas, H.J.D. and Bjorkman, A.D. (2019) Plant traits inform predictions of tundra responses to global change. *New Phytologist,* 221 (4), pp. 1742-1748.

Myers-Smith, I., Kerby, J.T., Phoenix, G.K., Bjerke, J.W., Epstein, H.E., Assmann, J.J., John, C., Andreu-Hayles, L., Angers-Blondin, S. and Beck, P.S.A. (2020) Complexity revealed in the greening of the Arctic. *Nature Climate Change,* 10 (2), pp. 106-117.

Myneni, R.B., Keeling, C.D., Tucker, C.J., Asrar, G. and Nemani, R.R. (1997) Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature,* 386 (6626), pp. 698-702.

Nadelhoffer, K.J., Giblin, A.E., Shaver, G.R. and Laundre, J.A. (1991) Effects of temperature and substrate quality on element mineralization in six arctic soils. *Ecology,* 72 (1), pp. 242-253.

Nauta, A.L., Heijmans, M.M.P.D., Blok, D., Limpens, J., Elberling, B., Gallagher, A., Li, B., Petrov, R.E., Maximov, T.C. and Van Huissteden, J. (2015) Permafrost collapse after shrub removal shifts tundra ecosystem to a methane source. *Nature Climate Change,* 5 (1), pp. 67-70.

Nieke J and Rast M (2018) Towards the copernicus hyperspectral imaging mission for the environment (CHIME) *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)* 157–9
Oberbauer, S.F., Elmendorf, S.C., Troxler, T.G., Hollister, R.D., Rocha, A.V., Bret-Harte, M., Dawes, M.A., Fosaa, A.M., Henry, G. and Høye, T.T. (2013) Phenological response of tundra plants to background climate variation tested using the International Tundra Experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 368 (1624), pp. 20120481.
Oldeland, J., Große-Stoltenberg, A., Naftal, L. and Strohbach, B.J. (2017) The potential of UAV derived image features for discriminating savannah tree species. In *The roles of remote sensing in nature conservation.* Springer, pp. 183-201.
Opedal, Ø.H., Armbruster, W.S. and Graae, B.J. (2015) Linking small-scale topography with microclimate, plant species diversity and intra-specific trait variation in an alpine landscape. *Plant Ecology & Diversity,* 8 (3), pp. 305-315.
Palace, M., Herrick, C., DelGreco, J., Finnell, D., Garnello, A.J., McCalley, C., McArthur, K., Sullivan, F. and Varner, R.K. (2018) Determining subarctic peatland vegetation using an unmanned aerial system (UAS). *Remote Sensing,* 10 (9), pp. 1498.
Park, T., Ganguly, S., Tømmervik, H., Euskirchen, E.S., Høgda, K., Karlsen, S.R., Brovkin, V., Nemani, R.R. and Myneni, R.B. (2016) Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data. *Environmental Research Letters,* 11 (8), pp. 084001.
Pattison, R.R., Jorgenson, J.C., Raynolds, M.K. and Welker, J.M. (2015) Trends in NDVI and tundra community composition in the Arctic of NE Alaska between 1984 and 2009. *Ecosystems,* 18 (4), pp. 707-719.
Pearson, R.G., Phillips, S.J., Loranty, M.M., Beck, P.S.A., Damoulas, T., Knight, S.J. and Goetz, S.J. (2013) Shifts in Arctic vegetation and associated feedbacks under climate change. *Nature climate change,* 3 (7), pp. 673-677.
Pignatti, S., Palombo, A., Pascucci, S., Romano, F., Santini, F., Simoniello, T., Umberto, A., Vincenzo, C., Acito, N. and Diani, M. (2013) The PRISMA hyperspectral mission: Science activities and opportunities for agriculture and land monitoring. IEEE, pp. 4558.
Pla, M., Duane, A. and Brotons, L. (2017) Potencial de las imágenes UAV como datos de verdad terreno para la clasificación de la severidad de quema de imágenes Landsat: aproximaciones a un producto útil para la gestión post incendio. *Revista de teledetección,* (49), pp. 91-102.
Pla, M., Bota, G., Duane, A., Balagué, J., Curcó, A., Gutiérrez, R. and Brotons, L. (2019) Calibrating Sentinel-2 Imagery with Multispectral UAV Derived Information to Quantify Damages in Mediterranean Rice Crops Caused by Western Swamphen (Porphyrio porphyrio). *Drones,* 3 (2), pp. 45.
Porder, S., Asner, G.P. and Vitousek, P.M. (2005) Ground-based and remotely sensed nutrient availability across a tropical landscape. *Proceedings of the National Academy of Sciences,* 102 (31), pp. 10909-10912.
Räsänen, A. and Virtanen, T. (2019) Data and resolution requirements in mapping vegetation in spatially heterogeneous landscapes. *Remote Sensing of Environment,* 230, pp. 111207.
Raynolds, M.K., Walker, D.A., Balser, A., Bay, C., Campbell, M., Cherosov, M.M., Daniëls, F.,J.A., Eidesen, P.B., Ermokhina, K.A. and Frost, G.V. (2019) A raster version

of the Circumpolar Arctic Vegetation Map (CAVM). *Remote Sensing of Environment,* 232, pp. 111297.
Reich P B, Walters M B and Ellsworth D S (1997) From tropics to tundra: global convergence in plant functioning *Proc. Natl Acad. Sci.* 94 13730–4
Riihimäki, H., Luoto, M. and Heiskanen, J. (2019) Estimating fractional cover of tundra vegetation at multiple scales using unmanned aerial systems and optical satellite data. *Remote Sensing of Environment,* 224, pp. 119-132.
Robinson, S.A., Turnbull, J.D. and Lovelock, C.E. (2005) Impact of changes in natural ultraviolet radiation on pigment composition, physiological and morphological characteristics of the Antarctic moss, Grimmia antarctici. *Global Change Biology,* 11 (3), pp. 476-489.
Roth, K.L., Roberts, D.A., Dennison, P.E., Alonzo, M., Peterson, S.H. and Beland, M. (2015) Differentiating plant species within and across diverse ecosystems with imaging spectroscopy. *Remote Sensing of Environment,* 167, pp. 135-151.
Royles J, Amesbury M J, Convey P, Griffiths H, Hodgson D A, Leng M J and Charman D J (2013) Plants and soil microbes respond to recent warming on the Antarctic Peninsula *Curr. Biol.* 23 1702–6
Serbin, S.P., Wu, J., Ely, K.S., Kruger, E.L., Townsend, P.A., Meng, R., Wolfe, B.T., Chlus, A., Wang, Z. and Rogers, A. (2019) From the Arctic to the tropics: multibiome prediction of leaf mass per area using leaf reflectance. *New Phytologist,* 224 (4), pp. 1557-1568.
Serbin S P, Singh A, McNeil B E, Kingdon C C and Townsend P A (2014) Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species *Ecological Applications* 24 1651–69
Shipley B, Lechowicz M J, Wright I and Reich P B (2006) Fundamental trade-offs generating the worldwide leaf economics spectrum *Ecology* 87 535–41
Siewert, M.B. and Olofsson, J. (2020) Scale-dependency of Arctic ecosystem properties revealed by UAV. *Environmental Research Letters,* 15 (9), pp. 094030.
Sivertson A H, Stodle D, Solbo S and Storvold R A new airborne remote sensing platform at Svalbard *Poster Session Presented at Svalbard Science Conf.: Svalbard in a pan–Arctic Perspective (Oslo, Norway, 5–6 November 2019)*
Soudzilovskaia, N.A., Elumeeva, T.G., Onipchenko, V.G., Shidakov, I.I., Salpagarova, F.S., Khubiev, A.B., Tekeev, D.K. and Cornelissen, J.H.C. (2013) Functional traits predict relationship between plant abundance dynamic and long-term climate warming. *Proceedings of the National Academy of Sciences,* 110 (45), pp. 18180-18184.
Stocker, T.F., Qin, D., Plattner, G., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (2013) Climate change 2013: The physical science basis. *Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change,* 1535.
Stow, D., Nichol, C.J., Wade, T., Assmann, J.J., Simpson, G. and Helfter, C. (2019) Illumination geometry and flying height influence surface reflectance and NDVI derived from multispectral UAS imagery. *Drones,* 3 (3), pp. 55.
Sturm, M., Schimel, J., Michaelson, G., Welker, J.M., Oberbauer, S.F., Liston, G.E., Fahnestock, J. and Romanovsky, V.E. (2005) Winter biological processes could help convert arctic tundra to shrubland. *Bioscience,* 55 (1), pp. 17-26.

Suvanto, S., Le Roux, P.C. and Luoto, M. (2014) Arctic-alpine vegetation biomass is driven by fine-scale abiotic heterogeneity. *Geografiska Annaler: Series A, Physical Geography,* 96 (4), pp. 549-560.
Tape, K.D., Christie, K., Carroll, G. and O'Donnell, J.A. (2016) Novel wildlife in the A rctic: the influence of changing riparian ecosystems and shrub habitat expansion on snowshoe hares. *Global Change Biology,* 22 (1), pp. 208-219.
Theodose, T.A. and Bowman, W.D. (1997) The influence of interspecific competition on the distribution of an alpine graminoid: evidence for the importance of plant competition in an extreme environment. *Oikos,* , pp. 101-114.
Thomas, H.J.D., Myers-Smith, I.H., Bjorkman, A.D., Elmendorf, S.C., Blok, D., Cornelissen, J.H.C., Forbes, B.C., Hollister, R.D., Normand, S. and Prevéy, J.S. (2019) Traditional plant functional groups explain variation in economic but not size-related traits across the tundra biome. *Global Ecology and Biogeography,* 28 (2), pp. 78-95.
Thomson E R, Malhi Y, Bartholomeus H, Oliveras I, Gvozdevaite A, Peprah T, Suomalainen J, Quansah J, Seidu J and Adonteng C (2018) Mapping the leaf economic spectrum across West African tropical forests using UAV-acquired hyperspectral imagery *Remote Sens.* 10 1532T
Tomassen, H.B.M., Smolders, A.J.P., Limpens, J., Lamers, L.P.M. and Roelofs, J.G.M. (2004) Expansion of invasive species on ombrotrophic bogs: desiccation or high N deposition? *Journal of Applied Ecology,* 41 (1), pp. 139-150.
Tømmervik, H., Karlsen, S., Nilsen, L., Johansen, B., Storvold, R., Zmarz, A., Beck, P.S., Høgda, K.A., Goetz, S. and Park, T. (2014) Use of unmanned aircraft systems (UAS) in a multi-scale vegetation index study of arctic plant communities in Adventdalen on Svalbard.
Turetsky, M.R., Bond-Lamberty, B., Euskirchen, E., Talbot, J., Frolking, S., McGuire, A.D. and Tuittila, E. (2012) The resilience and functional role of moss in boreal and arctic ecosystems. *New Phytologist,* 196 (1), pp. 49-67.
Turetsky, M.R., Mack, M.C., Hollingsworth, T.N. and Harden, J.W. (2010) The role of mosses in ecosystem succession and function in Alaska's boreal forest. *Canadian Journal of Forest Research,* 40 (7), pp. 1237-1264.
Turner, D., Lucieer, A., Malenovský, Z., King, D. and Robinson, S.A. (2018) Assessment of Antarctic moss health from multi-sensor UAS imagery with Random Forest Modelling. *International journal of applied earth observation and geoinformation,* 68, pp. 168-179.
Ullah, S., Skidmore, A.K., Ramoelo, A., Groen, T.A., Naeem, M. and Ali, A. (2014) Retrieval of leaf water content spanning the visible to thermal infrared spectra. *ISPRS journal of photogrammetry and remote sensing,* 93, pp. 56-64.
Van der Wal, R. and Brooker, R.W. (2004) Mosses mediate grazer impacts on grass abundance in arctic ecosystems. *Functional Ecology,* 18 (1), pp. 77-86.
Van Leeuwen, W. and Huete, A.R. (1996) Effects of standing litter on the biophysical interpretation of plant canopies with spectral indices. *Remote Sensing of Environment,* 55 (2), pp. 123-138.
Vanderpuye, A.W., Elvebakk, A. and Nilsen, L. (2002) Plant communities along environmental gradients of high-arctic mires in Sassendalen, Svalbard. *Journal of Vegetation Science,* 13 (6), pp. 875-884.

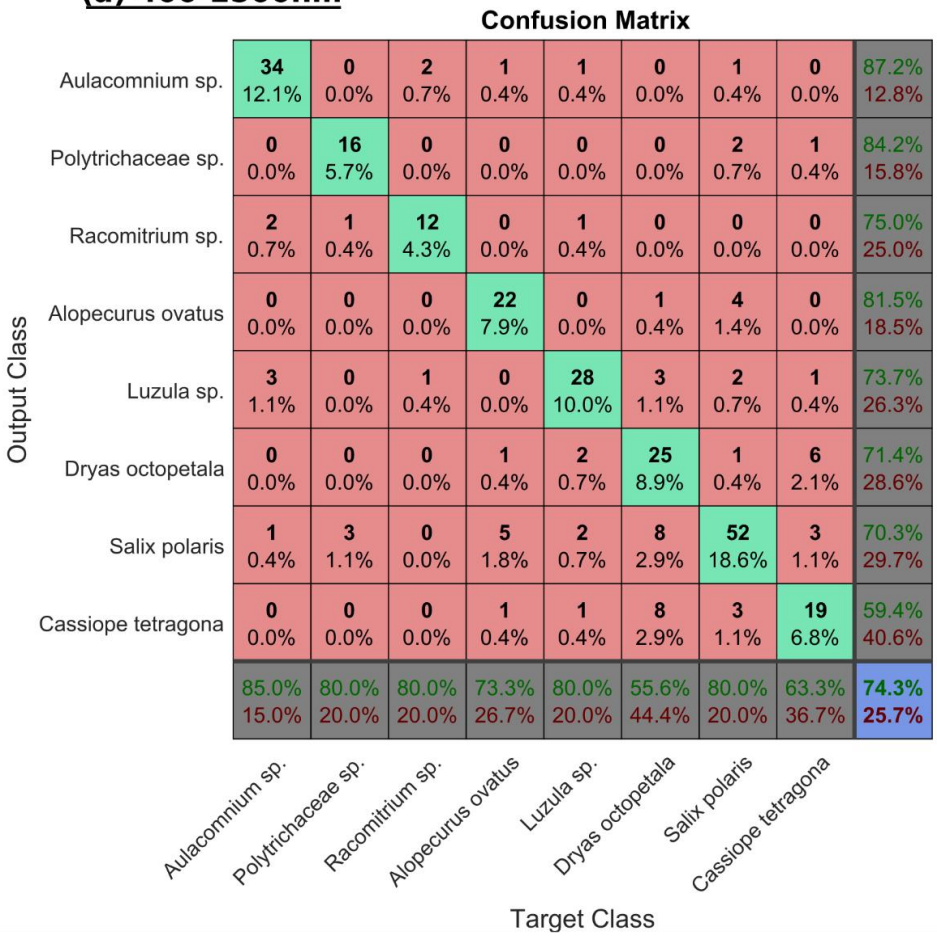
Wahren, C., Walker, M.D. and Bret-Harte, M.S. (2005) Vegetation responses in Alaskan arctic tundra after 8 years of a summer warming and winter snow manipulation experiment. *Global Change Biology,* 11 (4), pp. 537-552.
Wookey, P.A., Aerts, R., Bardgett, R.D., Baptist, F., Bråthen, K.A., Cornelissen, J.H.C., Gough, L., Hartley, I.P., Hopkins, D.W. and Lavorel, S. (2009) Ecosystem feedbacks and cascade processes: understanding their role in the responses of Arctic and alpine ecosystems to environmental change. *Global Change Biology,* 15 (5), pp. 1153-1172.
Wright I J, Reich P B, Westoby M, Ackerly D D, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen J H C and Diemer M (2004) The worldwide leaf economics spectrum *Nature* **428** 821–7
Zhang, Y., Song, C., Band, L.E., Sun, G. and Li, J. (2017) Reanalysis of global terrestrial vegetation trends from MODIS products: Browning or greening? *Remote Sensing of Environment,* 191, pp. 145-155.
Zwolicki, A., Zmudczyńska-Skarbek, K.M., Iliszko, L. and Stempniewicz, L. (2013) Guano deposition and nutrient enrichment in the vicinity of planktivorous and piscivorous seabird colonies in Spitsbergen. *Polar Biology,* 36 (3), pp. 363-372.

5.13: Supplementary material

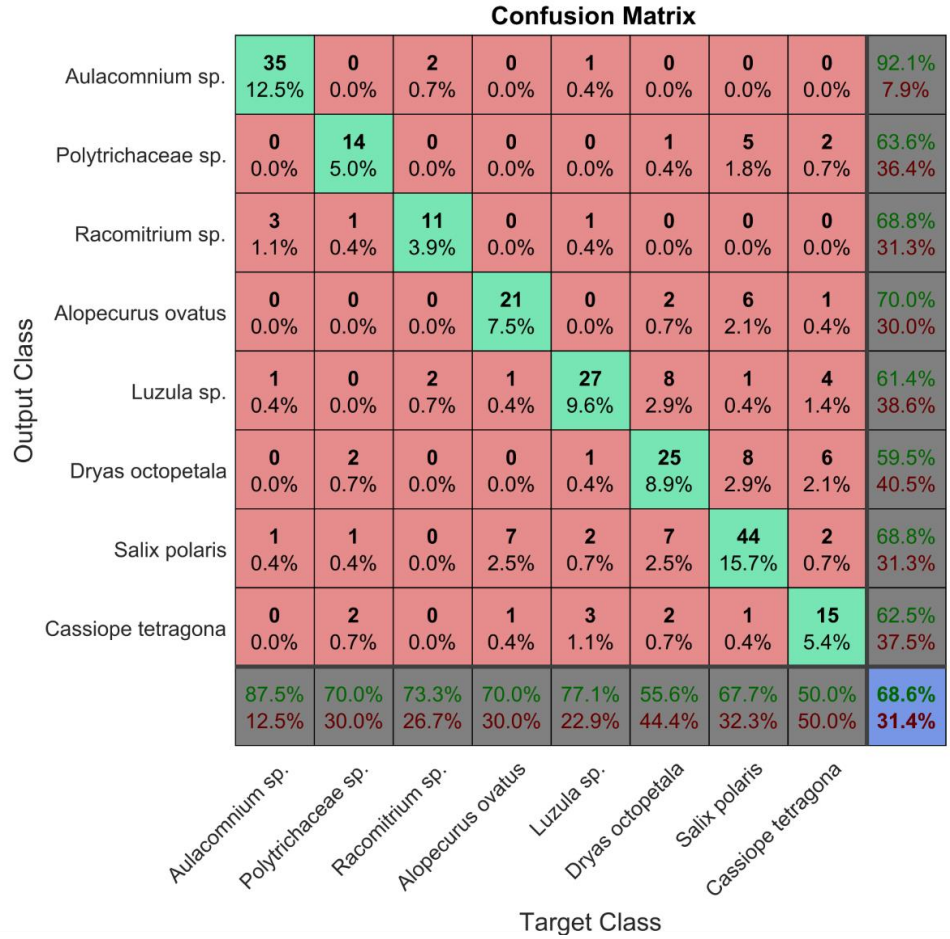
Table S5.1: Average plant trait values ± 1 standard deviation for turf samples collected at the Birdcliffs and ITEX site. Three trait measurements were taken per turf. Average values are given for each plant functional type (moss, graminoids and dwarf shrubs), with separate values for the senesced shrub samples. No senesced samples were collected at the Birdcliffs. *n* = number of measurements.

Birdcliffs					
Trait	**Moss**	**Graminoids**	**Dwarf shrubs**	**Senesced shrubs**	**Average**
n	45	42	48	-	-
Water (%)	70 ± 8.9	70 ± 9.5	56 ± 9.5	-	65
Nitrogen (%)	0.91 ± 0.46	2.08 ± 0.84	1.92 ± 0.68	-	1.64
Phosphorus (%)	0.08 ± 0.05	0.20 ± 0.11	0.16 ± 0.06	-	0.15
C:N ratio	45 ± 21	24 ± 14	27 ± 10	-	32
ITEX					
Trait	**Moss**	**Graminoids**	**Dwarf shrubs**	**Senesced shrubs**	**Average**
n	15	3	30	18	-
Water (%)	67 ± 12	60 ± 2.6	50 ± 12	34 ± 8	53
Nitrogen (%)	0.90 ± 0.45	1.78 ± 0.09	1.56 ± 0.71	0.86 ± 0.35	1.23
Phosphorus (%)	0.08 ± 0.07	0.08 ± 0.01	0.22 ± 0.19	0.08 ± 0.05	0.15
C:N ratio	88 ± 38	33 ± 1.71	35 ± 9	44 ± 12	50

(a) 400-2500nm



(b) 400-1000nm



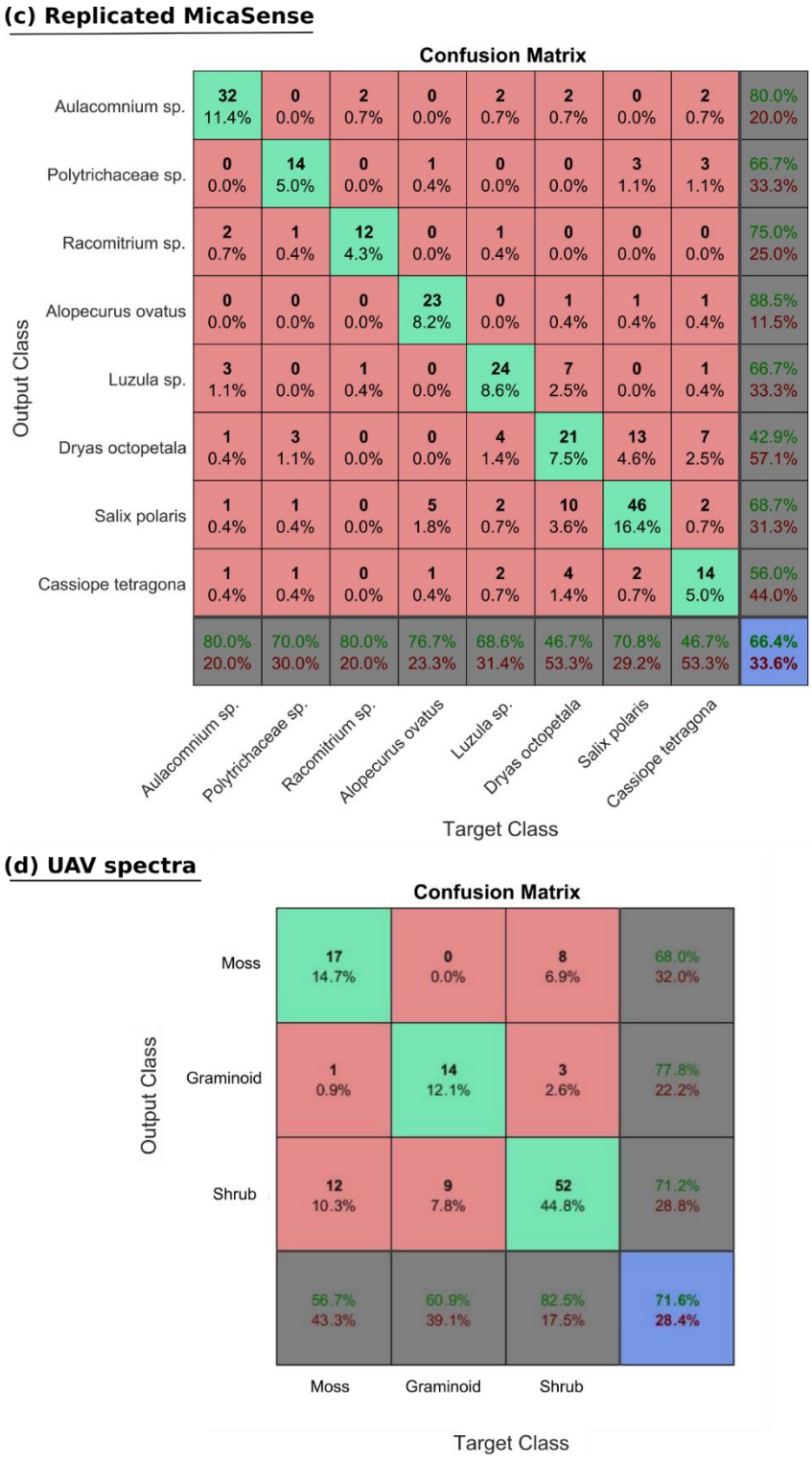


Figure S5.1 (a-c): Confusion matrices for the Random Forest species classification in *figure 4*. Produced using field spectroscopy data. The spectra were classified three times: (a) using the full hyperspectral range from 400-2500 nm, (b) just the VIS-NIR region from 400-1000 nm and (c) downscaled spectra to match the five bands represented by the MicaSense camera (blue, green, red, rededge and near-infrared). (d) Confusion matrix corresponding to the Random Forest plant functional type classification in *figure 6.* Produced using UAV-collected reflectance spectra.

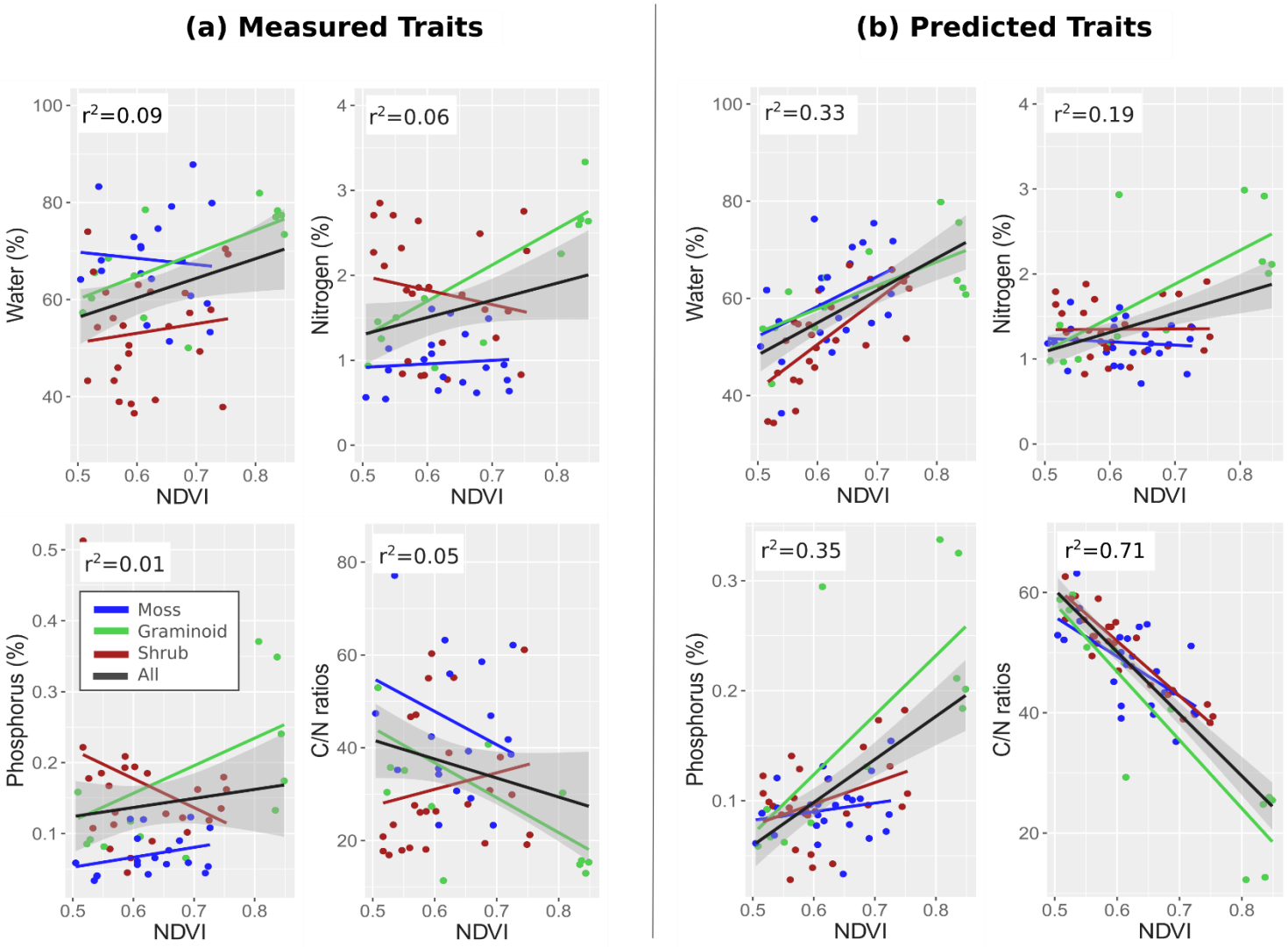


Figure S5.3: (a) Relationships between field-measured plant trait values and NDVI over the Birdcliff site. NDVI was calculated using UAV imagery. (b) the same as (a) but using predicted plant trait values based on UAV-PLSR models. r^2 values are presented for all the data (moss, graminoids and dwarf shrubs combined).

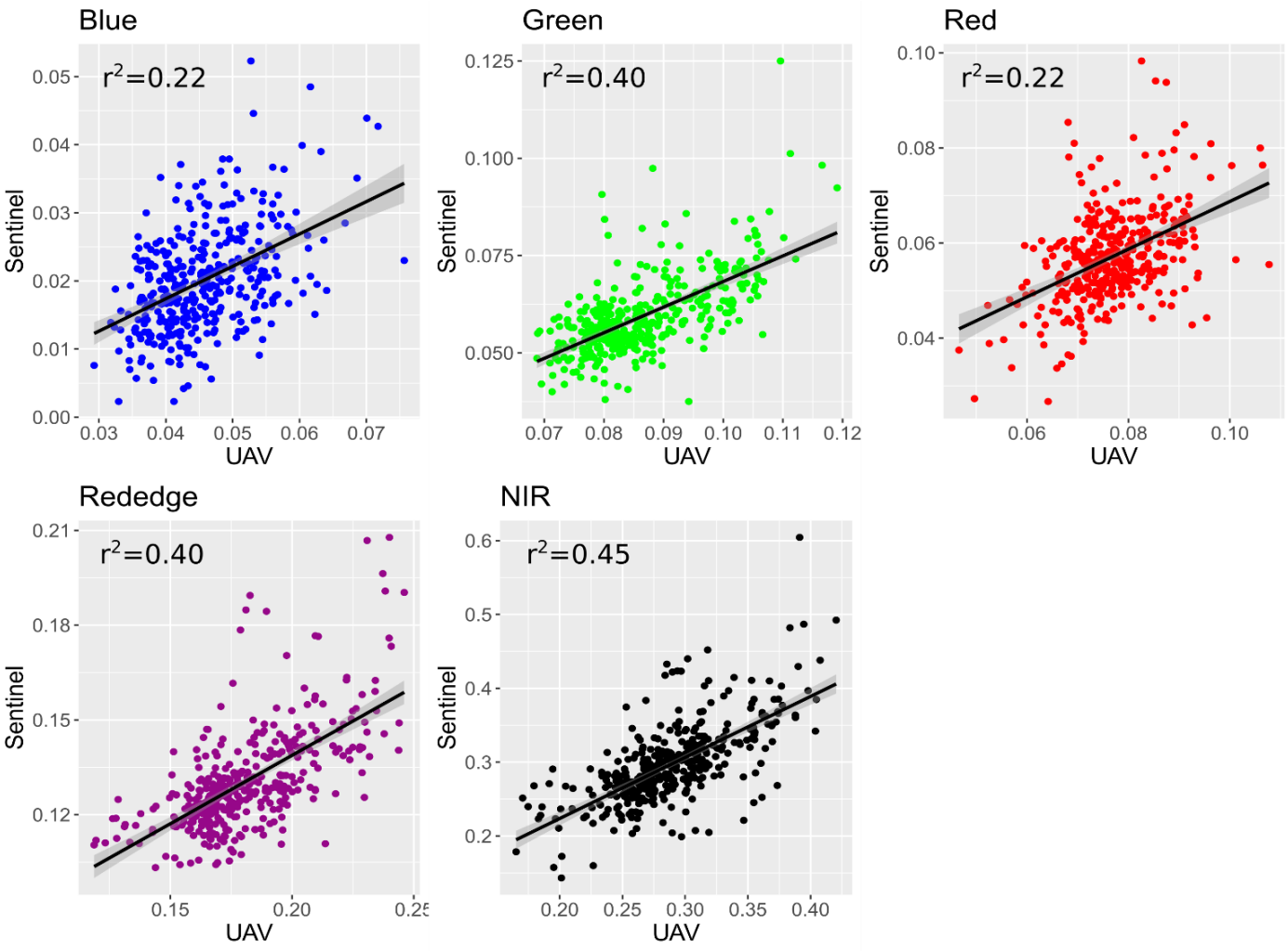


Figure S5.4: Relationships between 5-band UAV reflectance and Sentinel-2 surface reflectance (Bands 2,3,4,5 and 8) over the Birdcliff site. UAV reflectance was resampled using bilinear interpolation to the resolution and gridlines of the Sentinel imagery (10 m). Only vegetation pixels were used in the analysis, as classified by the Sentinel-2 Scene Classification Layer (SCL). The Sentinel image was taken one week after the UAV-collected imagery.

Chapter 6:

Mapping the Leaf Economic Spectrum across West African Tropical Forests Using UAV-Acquired Hyperspectral Imagery

Authors: Eleanor R. Thomson[1,2], Yadvinder Malhi[1], Harm Bartholomeus[3], Imma Oliveras[1], Agne Gvozdevaite[1], Theresa Peprah[4], Juha Suomalainen[3], John Quansah[4], John Seidu[4], Christian Adonteng[4], Andrew J. Abraham[2], Martin Herold[3], Stephen Adu-Bredu[4] and Christopher E. Doughty[2]

Affiliations:

[1] Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, UK

[2] School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011, USA

[3] Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands

[4] Forestry Research Institute of Ghana, Council for Scientific and Industrial Research, P.O. Box UP 63 KNUST, Kumasi, Ghana

[5] Finnish Geospatial Research Institute, National Land Survey of Finland, Geodeetinrinne 2, 02430 Masala, Finland

6.1. Chapter Statement

This chapter uses UAV-collected hyperspectral imagery to map biochemical leaf traits across a tropical forest in Ghana. Similarly to **Chapter 5**, it starts by using field spectroscopy to quantify biochemical traits at the leaf level (LMA, N, P, K, Ca, Mg). Building on the results in **Chapter 5**, it tests methods of trait retrieval using UAV-collected hyperspectral imagery, across an extremely species rich ecosystem, with the added complexity of canopy structure, variable leaf angle and shadowing. It links the resulting biochemical trait maps with the leaf economic spectrum to draw inferences about wider ecosystem functioning.

This chapter was published on 24th September, 2018 in *Remote Sensing* (https://doi.org/10.3390/rs10101532). Results were presented at the Association of Tropical Biology and Conservation conference in Kuching, Malaysian Borneo.

6.2 Author Contributions

The ideas, research questions, study design, data analysis and writing presented in this chapter are the work of the first author, Eleanor Thomson. These activities greatly benefited from the contributions of Chris Doughty and Yadvinder Malhi.

Leaf trait collection was carried out by Imma Oliveras, Agne Gvozdevaite, Theresa Peprah, John Quansah, John Seidu, Christian Adonteng and Stephen Adu-Bredu. Field spectroscopy data was collected by Chris Doughty. UAV imagery collection was carried out by Harm Bartholomeus, Martin Herold and Juha Suomalainen from Wageningen University and Research. Chemical analysis was undertaken at Wageningen University and Research.

Logistical and financial support for the research was provided by Stephen Adu-Bredu, Yadvinder Malhi and Chris Doughty.

All authors provided constructive comments during the writing of the manuscript.

6.3 Abstract

The leaf economic spectrum (LES) describes a set of universal trade-offs between leaf mass per area (LMA), leaf nitrogen (N), leaf phosphorus (P) and leaf photosynthesis that influence patterns of primary productivity and nutrient cycling. Many questions regarding vegetation-climate feedbacks can be addressed with a better understanding of LES traits and their controls. Remote sensing offers enormous potential for generating large-scale LES trait data. Yet so far, canopy studies have been limited to imaging spectrometers onboard aircraft, which are rare, expensive to deploy and lack fine-scale resolution. In this study, we measured VNIR (visible-near infrared (400–1050 nm)) reflectance of individual sun and shade leaves in 7 one-ha tropical forest plots located along a 1200–2000 mm precipitation gradient in West Africa. We collected hyperspectral imaging data from 3 of the 7 plots, using an octocopter-based unmanned aerial vehicle (UAV), mounted with a hyperspectral mapping system (450–950 nm, 9 nm FWHM). Using partial least squares regression (PLSR), we found that the spectra of individual sun leaves demonstrated significant ($p < 0.01$) correlations with LMA and leaf chemical traits: r^2 = 0.42 (LMA), r^2 = 0.43 (N), r^2 = 0.21 (P), r^2 = 0.20 (leaf potassium (K)), r^2 = 0.23 (leaf calcium (Ca)) and r^2 = 0.14 (leaf magnesium (Mg)). Shade leaf spectra displayed stronger relationships with all leaf traits. At the airborne level, four of the six leaf traits demonstrated weak ($p < 0.10$) correlations with the UAV-collected spectra of 58 tree crowns: r^2 = 0.25 (LMA), r^2 = 0.22 (N), r^2 = 0.22 (P), and r^2 = 0.25 (Ca). From the airborne imaging data, we used LMA, N and P values to map the LES across the three plots, revealing precipitation and substrate as co-dominant drivers of trait distributions and relationships. Positive N-P correlations and LMA-P anticorrelations followed typical LES theory, but we found no classic trade-offs between LMA and N. Overall, this study demonstrates the application of UAVs to generating LES information and advancing the study and monitoring tropical forest functional diversity.

Keywords: leaf traits; leaf economic spectrum; UAV; hyperspectral; spectroscopy; tropical forest; PLSR; Ghana; West Africa

6.4. Introduction

Plant functional traits refer to specific features about a plant's morphological and physiological characteristics that determine factors such as its preferred environment, growth rate, life strategy, dispersal ability and tolerance of pests and hazards (Kattge *et al.* 2011). As plant traits tend to represent fundamental trade-offs that determine a species' ecological performance, there is currently much interest in using plant traits to link species richness to functional diversity to better understand the influence of biological communities on ecosystem function (Díaz *et al.* 2004, 2016; McGill *et al.* 2006; Lavorel *et al.* 2011; Enquist *et al.* 2015) and improve predictions of ecosystem response to environmental change (Webb *et al.* 2010; Drenovsky *et al.* 2012; Lavorel *et al.* 2013).

Leaf traits demonstrate a direct link to ecosystem function via the Leaf Economic Spectrum (LES). The LES describes an axis of six main leaf traits (leaf mass per area (LMA), leaf nitrogen (N), leaf phosphorus (P), leaf longevity, photosynthetic rate, and dark respiration rate) that covary in predictable ways and underpin plant resource acquisition–defence tradeoffs (Reich *et al*. 1997; Wright *et al*. 2004). On one end of the spectrum, species opt for a quick return on investments by producing 'cheap', short-lived leaves with high N and P concentrations that facilitate high photosynthesis rates and fast growth. These species are commonly found in high-fertility environments where short lived leaves from abscission and herbivory maintain relatively fast rates of nutrient cycling (Santiago, 2007; Cornwell *et al.* 2008). On the other end of the spectrum, species located in lower-precipitation and -nutrient regimes tend to trade rapid growth for persistence and invest in tough, well-protected leaves, with higher LMA values and lower N and P concentrations. They produce leaves that can better withstand herbivory and physical hazards, leading to increased leaf longevity and a lower rate of nutrient cycling (Hoffmann *et al*. 2005; Poorter and Markesteijn 2008). The LES has been shown to be a good indicator of plant performance (Poorter and Bongers, 2006; Sterck *et al.* 2006), with a species' position on the spectrum describing the plant's overall strategy (quick growing but ephemeral versus slow growing but enduring). Importantly, this spectrum has been shown to operate similarly across a wide variety of species and biomes, with important implications for global dynamic vegetation and climate models (Sakschewski *et al*. 2015).

Despite the fact that functional traits have been found to form the basis of a more quantitative and predictive ecology, a major limitation to this research has been a lack of spatially extensive trait data (Jetz *et al.* 2016). This lack of trait data is particularly severe in tropical forests (Asner *et al*. 2015). Tall trees, high species diversity and the labour-intensive nature of collecting trait data, combined with the inaccessibility of many tropical regions, means measuring and monitoring functional traits in tropical areas is a time-consuming and expensive process. Yet tropical forests comprise 40–50% of the total carbon in terrestrial vegetation (Dixon *et al.* 1994; Pan *et al.* 2011; Ciais *et al.* 2013), store a third of global soil carbon (Jobbágy *et al*. 200) and currently absorb around 18% of human CO_2 emissions (Lewis *et al.* 2009). Thus, determining the functional diversity of tropical forest canopies is critical to predicting future climate–biosphere interactions, as a small perturbation in the net balance of tropical forest carbon cycling would have significant climate implications.

Remote sensing offers enormous potential for generating large-scale trait data. Using high-resolution spectroscopy (VNIR (400–1100) or VSWIR (400–2500 nm)) to measure plant foliar traits has a long scientific record (Gates *et al*. 1965; Knipling, 1970). Specific spectral absorption features have been directly linked to a leaf's concentration of cellulose, lignin, chlorophylls, N, starch, oil, proteins and water (see Curran (1989) for a well-known synthesis). For example, the measurement of N, bound in chlorophyll and proteins, is associated with the bandwidths between 460–480, 650–670 and 1510–1600 nm (Kokaly *et al*. 2001; Smith *et al*. 2003), while LMA can be indirectly estimated using variations in leaf thickness and leaf water content expressed across the 700–1300 nm range (Asner *et al*. 2008; Asner *et al*. 2011a; Doughty *et al*. 2011).

In the last decade, significant headway has been made in the remote sensing of tropical forest leaf traits, both at the individual leaf level and at the canopy level (Asner *et al.* 2016a). Relationships between spectra and foliar traits have been developed using partial least squares regression analysis (PLSR). The PLSR technique is useful because it does not focus on a small number of wavelengths but uses the entire leaf spectral continuum in its analysis. PLSR is designed to handle situations where the number of predictor variables (e.g. 101 spectral bands) is higher than the number of observations, avoiding earlier problems of collinearity and statistical overfitting (Wold *et al.* 1984).

Using PLSR, relationships have also been found between leaf spectra and macronutrients, such as P, K, Ca and Mg, that have no direct expression in the VNIR or VSWIR, but often demonstrate correlations with other traits (Asner *et al.* 2015; Asner *et al*. 2011b; Asner *et al.* 2012a; Gökkaya *et al.* 2015; Stein *et al*. 2014).

Similarly, using PLSR and spectroscopy, it has been found that 'higher-order' foliar traits, such as photosynthesis and respiration, can be predicted with a high degree of accuracy and precision (Doughty *et al.* 2011). Globally, the largest fraction of net primary productivity (NPP: 30–40%) is thought to occur in tropical forests (Cleveland *et al*. 2013; Grace *et al*. 2001; Malhi and Grace, 2000). Yet there are fears that, in the Amazon at least, this value may be decreasing (Brienen *et al*. 2015). Satellite-based techniques for measuring tropical NPP have attracted recent criticism, as they demonstrate little spatial or temporal agreement with plot-based values (Cleveland *et al*. 2015). Thus, UAV-based spectroscopy may offer a new approach for estimating NPP across small-scale tropical forest areas.

This paper aims to investigate the use of UAVs for the mapping and monitoring of tropical forest canopy traits. Using a similar method to leaf spectroscopy, airborne hyperspectral imaging spectrometers measure solar radiation reflected from the forest canopy at numerous narrow wavelengths, producing estimates of canopy foliar traits at scales unachievable via field and lab approaches (Huang *et al*. 2009; Cho *et al.* 2012; Asner *et al.* 2016b). In the last ten years, developments and improvements in UAVs, together with the miniaturization of sensing technologies, has opened up previously unavailable opportunities for remote sensing. In comparison to aircraft, UAVs are accessible to a wide array of researchers, they can operate at short notice without a planned schedule, and provide high spatial and spectral resolution data at low cost. Although their flight duration is much shorter than that of an aircraft, limiting their area of data collection, they can fly slower, at much lower altitudes, and thus acquire higher spatial and temporal resolution data (Pajares *et al.* 2015). Using UAVs to map and monitor tropical forest canopy traits would have distinct advantages over aircraft in that they would (i) enable a wider proportion of the research community to quantify canopy traits, increasing spatial coverage and innovation in this area, (ii) facilitate the swift upscaling of field study results to the immediate geographic area, (iii) be able to be deployed quickly after a disturbance

event, such as a drought or fire, to measure its effects, (iv) be able to be deployed on multiple occasions over a short time period to monitor temporal changes, and (v) assist in a wide range of conservation related projects and programmes and promote an interest in plant traits outside of the academic community (Asner and Martin, 2008; Asner *et al*. 2017).

Using field-measured traits and spectral data collected across a tropical forest precipitation gradient in West Africa, this study expects:

i) High-resolution VNIR spectroscopy will be able to predict LMA, N, P, K, Ca and Mg in sun and shade leaves at the individual leaf level. While sun leaves have received a lot of attention in remote sensing studies due to their natural prominence in airborne imaging data, shade leaves (those located under the canopy in the understory) constitute the majority of tropical forest foliage and are known to have a strong influence on canopy spectral signatures in the near-infrared (NIR). We expect that VNIR spectroscopy will be able to predict traits directly expressed in leaf spectra (LMA and N) at the highest level of accuracy and traits indirectly in spectra (P, K and CA) at lower accuracy.

ii) The same foliar traits to be quantified with similar accuracy and precision at the tree crown level using UAV-collected hyperspectral imagery.

iii) The spatial distribution of traits and their relationships will subscribe to general LES theory.

6.5. Materials and Methods

The Study Sites

Data collection was carried out at three sites across a precipitation gradient in Ghana: Ankasa Conservation Area, Bobiri Forest Reserve and Kogyae Strict Nature Reserve (**Figure 6.1**). The three study sites form part of the Global Ecosystems Monitoring network (GEM: Malhi *et al*. 2021), which is a long-term effort to measure and monitor forests and their response to climate change.

Ankasa is located in south-west Ghana, in tropical wet evergreen forest and receives over 2000 mm of rain a year, spread over two rainy seasons—April to July and September to November. Soils are loamy, highly weathered, acidic, and high in aluminium (Moore *et al.* 2017). Data were taken from two 1-ha forest plots; Ankasa 01 located on mature upland soils and Ankasa 03 located in low riverein forest that is periodically inundated during the rainy season. See **Table 6.1** for more details.

Bobiri is located in moist semi-deciduous forest and receives an annual precipitation of approximately 1500 mm, spread over the same two rainy seasons described above. Soils in Bobiri are similarly loamy but less acidic than in Ankasa, and unusually high in base cations. Data were taken from two 1-ha forest plots (Bobiri 01 and Bobiri 02). Bobiri 01 is located within the Strict Nature Reserve with no record of logging. According to Forest Commission (FC) records, Bobiri 02 was last (very lightly) logged in 1959. This logging comprised of 1–2 stems per hectare, which was rigorously enforced by the FC.

Kogyae is located in the transitional forest zone, with an average rainfall of 1200–1300 mm. To reflect the full extent of the forest–savannah transition landscape, data were taken from three plots—Kogyae 02 located in dry forest, Kogyae 04 located in forest-savannah transition and Kogyae 05 located in woody savanna. Soils in Kogyae are sandy with low water and nutrient retention properties. Similarly to Bobiri, they are high in base cations, either due to local depositional features or Saharan dust deposits, transported during the Harmattan winds (Tiessen *et al*. 1991).

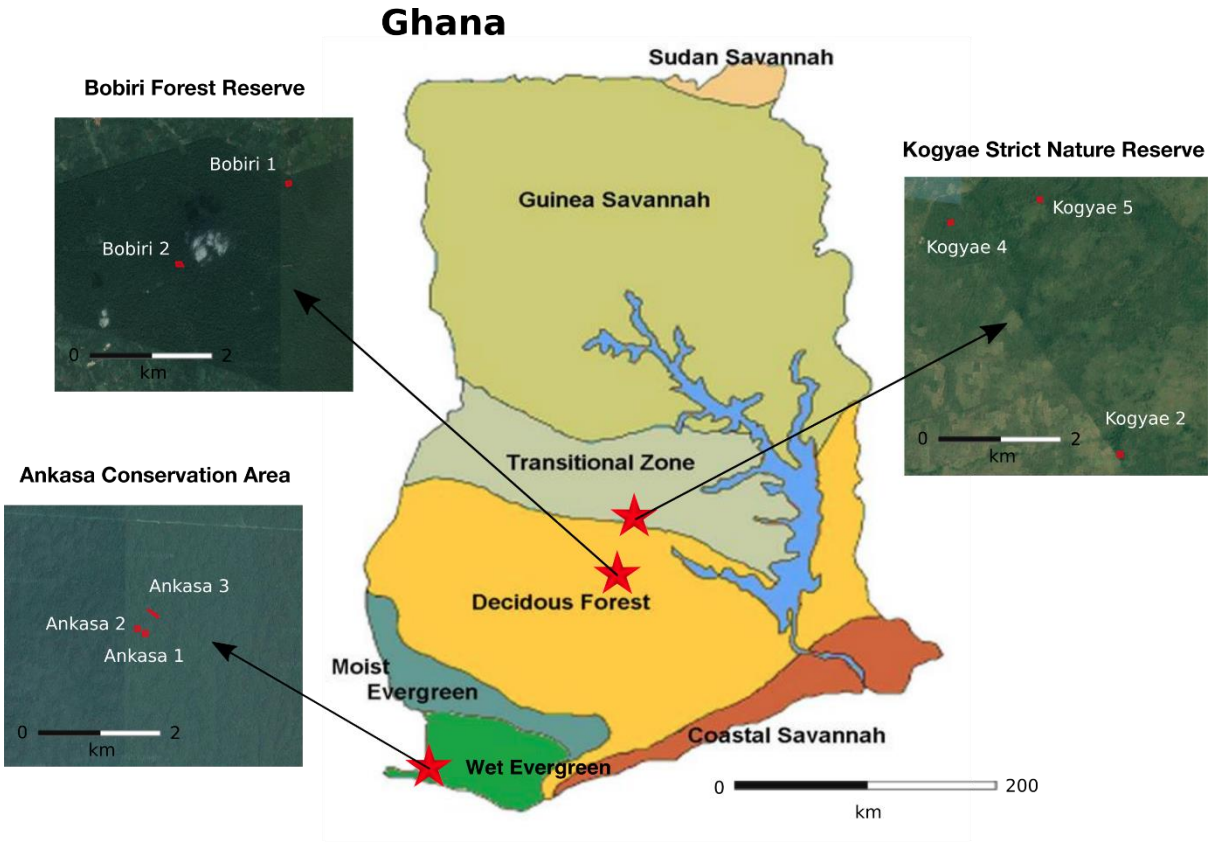


Figure 6.1. Map showing the major ecological zones of Ghana and location of research sites (Ankasa Conservation Area, Bobiri Forest Reserve and the Kogyae Strict Nature Reserve). Within each site, the one-ha study plots are outlined. Satellite photographs taken from Google Earth with permission.

	Ankasa Conservation Area	Bobiri Forest Reserve	Kogyae Strict Nature Reserve
Latitude (°)	5.2680	6.6910	7.3012
Longitude (°)	−2.6955	−1.3389	−1.1649
Elevation (m.a.s.l)	114	254	225
Mean annual air temperature (°C)	25	25.7	26.4
Mean annual precipitation (mm yr^{-1})	2050	1500	1200
Mean maximum climatological water deficit (mm)	−13	−374	−412
Soil pH	4.27	6.05	6.07
Soil N (%)	0.17	0.16	0.06
Soil C (%)	2.61	1.71	0.72
P_{total} (mg kg^{-1})	147	258	67
Ca_{ex}	1.34	32.81	18.91
K_{ex}	0.83	1.25	1.09
Mg_{ex}	3.45	11.00	6.22
Al_{ex}	18.44	0.89	0.02
Sand (%)	63	47	83
Clay (%)	22	29	2
Silt (%)	15	24	15
Leaf photosynthesis rate (μmol m^{-2} s^{-1})	5.87	7.75	7.74
Leaf residence time (months)	9–10	4–6	6.5–8
Aboveground coarse wood residence time (months)	99 ± 22.1	39. 65 ± 8.88	40.60 ± 9.08
NPP (Mg C ha^{-1} yr^{-1})	13.12 ± 0.79	11.74 ± 0.95	10.19 ± 0.78

Table 6.1. Meteorological, soil and productivity characteristics for the 3 tropical forest study sites located along a 1200–2000 mm rainfall gradient in Ghana. P_{total} = total soil phosphorus pool. Ca_{ex}, Mg_{ex}, K_{ex}, Na_{ex}, Al_{ex} = exchangeable calcium, magnesium, potassium and aluminum concentrations. ±1 standard deviation. Adapted from Moore *et al.* (2017). See Moore *et al.* (2017) and Gvozdevaite *et al.* (2018) for more details.

Leaf Trait Sampling

Fieldwork was carried out between October 2014 and March 2015. In Kogyae, leaf traits were sampled between October and November. In Ankasa and Bobiri, leaf traits were sampled between December and March. Leaf traits were collected as part of the KWAEEMMA traits campaign ('Children of the Forest' in the local Twi language), a collaboration between GEM and the Forestry Research Institute of Ghana. For each one-ha forest plot, the most abundant species were chosen that represented over >55% of the basal area of each plot. From each of these species, the three largest individuals were sampled (based on diameter at breast height (DBH)). From these trees, six leaves were harvested from branches cut down by a professional tree climber; three from a sun branch in the upper canopy located in direct sunlight and three from a branch underneath the main crown located in complete shade. Leaf spectroscopy and trait measurements were collected for all leaves. Across the seven plots, this collection resulted in leaf spectra and corresponding trait measurements for 753 sun and shade leaves respectively.

Less than 12 h after collection, leaves were scanned using a digital scanner (Canon LiDE 110) to determine their area and then oven dried at 72 °C for three days. LMA was derived by dividing the leaf's dry mass (g) by its fresh one-sided area (m^2). The dried leaves were then used for chemical analysis. Leaf chemistry measurements (N, P, K, Ca, Mg) were carried out at Wageningen University using a mass spectrometer (see Moore *et al.* (2017) for more details). All leaf chemical traits are reported on a dry mass basis.

Leaf Spectroscopy Measurements

Hemispheric leaf reflectance measurements were carried out immediately in the field after branch harvesting. Measurements were taken on the adaxial surface of the leaf, halfway between the petiole and leaf tip and at the mid-point between the main vein and leaf edge. Care was taken to avoid large primary or secondary veins. The spectra spanned 750 bands between 325–1075 nm and were collected using an ASD Fieldspec Handheld 2, with fibre optic cable, contact probe and an illumination collimator (Analytical Spectral Devices, Boulder, CO, USA). For each leaf, an average of 25 spectra were taken. Measurements were collected with 136 ms integration time per spectrum. The spectrometer was optimised and calibrated for dark current and white light after every three measurements. The spectra were trimmed at the far edges to remove noise, producing a final range of 400–1050 nm. To control for outliers, spectra lying more than two standard deviations away from the mean in the visible light (550 nm), or NIR (900 nm), were flagged and checked.

Airborne Spectroscopy

The airborne imaging data were collected in the first two weeks of March, 2016. Data were taken from both Ankasa plots, and Bobiri 02. No airborne data were taken from Kogyae, as it was late dry season and the trees were mostly leafless.

The airborne imaging data were collected using the WageningenUR HYMSY system—a lightweight hyperspectral mapping system mounted on an octocopter UAV (Suomalainen *et al.* 2014). The HYMSY is comprised of a custom made pushbroom spectrometer (450–950 nm, 9 nm FWHM, 25 lines/s, 328 px/line), a photogrammetric camera, and a miniature GPS Inertial Navigation System. Each plot was covered by between 2–4 flights. The raw hyperspectral data was processed to reflectance factors using the empirical line method (Smith *et al.* 1999) with the help of a grey reference panel placed on the forest floor near the UAV take-off position. Using the HYMSY processing chain, each flight produced an RGB orthomosaic at 3 cm resolution, a Digital Surface Model (DSM) at 20 cm resolution, and a hyperspectral datacube at 50 cm resolution. For more details on the processing methodology see Suomalainen *et al.* (2014).

Analysis of the hyperspectral datacube was carried out in Environment for Visualizing Images (ENVI), version 5.3 (Exelis Visual Information Solutions, Boulder, CO, USA). Flight lines over each plot were stitched together using the ENVI mosaicking tool. All pixels with an NDVI < 0.6 were removed to screen out non-vegetation (Weier *et al.* 2000). Shadows were removed by masking out all pixels with a maximum reflectance of less than 0.2 across all wavelengths (Zhang *et al.* 2015).

Tree crown spectra were manually extracted for each individual tree from which field spectroscopy and trait measurements had been taken, and that exhibited crowns that lay > 50% in direct sunlit (**Figure 6.2**). Tree crowns were identified using the RGB orthomosaic, which displayed the forest canopy at high enough resolution to visually distinguish between different trees crowns. The tree crowns recognizable on the RGB orthomosaic were manually matched to the field-sampled trees by using x,y coordinates for each tree, species identification, known crown dimensions, and tree height from the DSM. Overall, this method resulted in n = 58 for matched tree canopy spectra across all

three plots. Tree crowns were manually delineated and identified in accordance with Dalponte *et al*. (2015) who found that the manual identification of tree crowns had a higher accuracy rate than those automatically, or even semi-automatically, delineated and identified from hyperspectral data. Our method was closely aligned with Dalponte *et al*. (2015); Leckie *et al*. (2003); and Clark *et al.* (2005), who similarly used tree location, species type, tree height and crown diameter to aerially identify tree crowns.

An average canopy spectrum for each plot was extracted using geo-referenced plot corner coordinates and the ENVI region of interest tool.

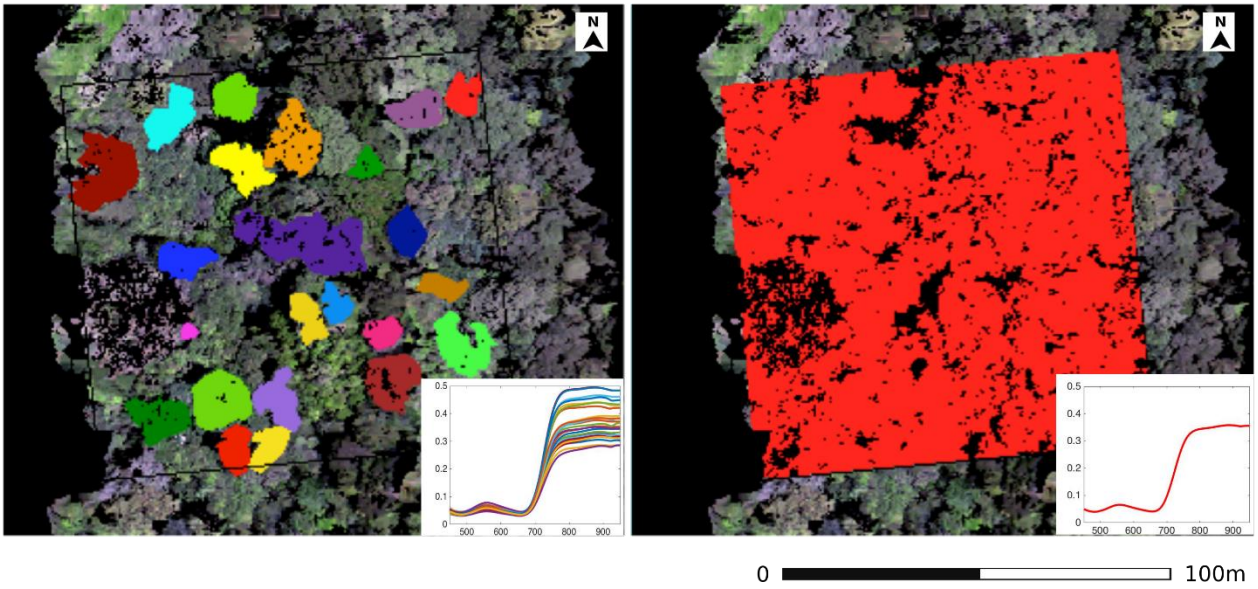


Figure 6.2: (Left) Tree crown spectra extracted from Ankasa 01 with NDVI < 0.6 and shadows masked. Spectra were only extracted for trees from which field data were taken and that exhibited crowns that lay > 50% in direct sunlight. (Right) Average plot spectrum extracted from Ankasa 01 with NDVI < 0.6 and shadows masked.

Spectra–Trait Analyses

To predict leaf traits using the individual sun and shade leaf spectra, we used PLSR analysis. The PLSR method is effective, as it uses the continuous spectrum as a single measurement, rather than carrying out a band-by-band analysis and reduces a large predictor matrix (650 spectral bands between 400–1050 nm) down to a few relatively uncorrelated latent factors (Haaland and Thomas, 1988). In addition, each part of the spectrum is weighted to determine its relative importance to predicting the foliar trait. We generated predictive models for each trait using the PLSregress function in MATLAB (MathWorks Inc., Natick, MA, USA). To create a completely independent validation dataset, the data was split 70:30 and the PLSR model was calibrated on a randomly selected 70% of the data, with the remaining 30% reserved as an independent testing set. Due to this random component, an average of 30 PLSR runs was taken for each spectra–trait analysis and the results averaged. To avoid overfitting the number of latent factors in each analysis, 30-fold cross validation was used on the calibration dataset. This process iteratively removes one sample from the input data set until the mean square error is minimized. The PLSR models were evaluated using four main metrics—r^2 for both the calibration (cal) and independent testing (val) datasets, root mean square error (RMSE), and RMSE as a percentage of the sample mean (%RMSE).

To quantify the airborne spectra–trait relationships, the 58 extracted tree crown spectra were matched with the field-collected sun leaf trait values and run through the same PLSR analysis described above.

Mapping Foliar Traits

To create maps of predicted canopy traits for each plot, the predictive models generated from the airborne spectra–trait PLSR analyses were applied to each pixel of the hyperspectral plot images using the following formula:

$$Y = b_0 + \sum_{i=1}^{101} b_i R(\lambda_i)$$

where Y is the trait value of interest, b_0 is the regression coefficient for the intercept computed from the PLSR airborne spectra–trait analyses, the b_i values are the regression coefficients for nanometers 450 through 950 nm at 5 nm intervals, and R is the reflectance value for each band (λ) of the UAV-collected hyperspectral data. Regression coefficients were taken from the 5 PLSR models that maximized the r^2 and minimized the RMSE value, and averaged. As models that maximized the r^2 did not always minimize the RMSE (and vice versa), it was found that an average of 5 models performed better than one model alone. A colour palette was added to the results. To our knowledge, this method generated the first maps of tropical forest canopy traits produced using UAV-collected spectra. Maps were only created for 4 out of the 6 leaf traits, as the r^2 and RMSE values for K and Mg were considered too weak to carry forward. The maps of predicted canopy LMA, N and P were combined using an RGB composite to generate a map of LES trait interactions.

To investigate the accuracy of the leaf trait maps, the average remotely sensed trait value for each plot was compared with the average field-collected value using r^2, RMSE and %RMSE. The field-collected trait values were weighted using a basal-weighted area and stem-weighted abundance method, as well as a simple mean (Asner *et al.* 2016c). To evaluate trait relationships, the remotely sensed trait values for all pixels across each plot were analysed using Pearson's correlation.

6.6. Results

Variation in Leaf Traits and Spectral Properties

Average leaf spectra across Ankasa, Bobiri and Kogyae (**Figure 6.3**) demonstrated a very similar shape, magnitude and range, with reflectance peaking in the NIR (700–1075 nm). The spectra showed little difference between sun and shade leaves (**Figure 6.3 (a-b)**). Across Ankasa 01, Ankasa 03 and Bobiri 02, the UAV-collected spectra demonstrated a similar shape to the individual leaf level spectra, but with significantly lower reflectance in the visible and NIR regions (**Figure 6.3(c–e)**). The variation in spectra was significantly higher for the UAV-collected spectra (**Figure 6.3(f)**).

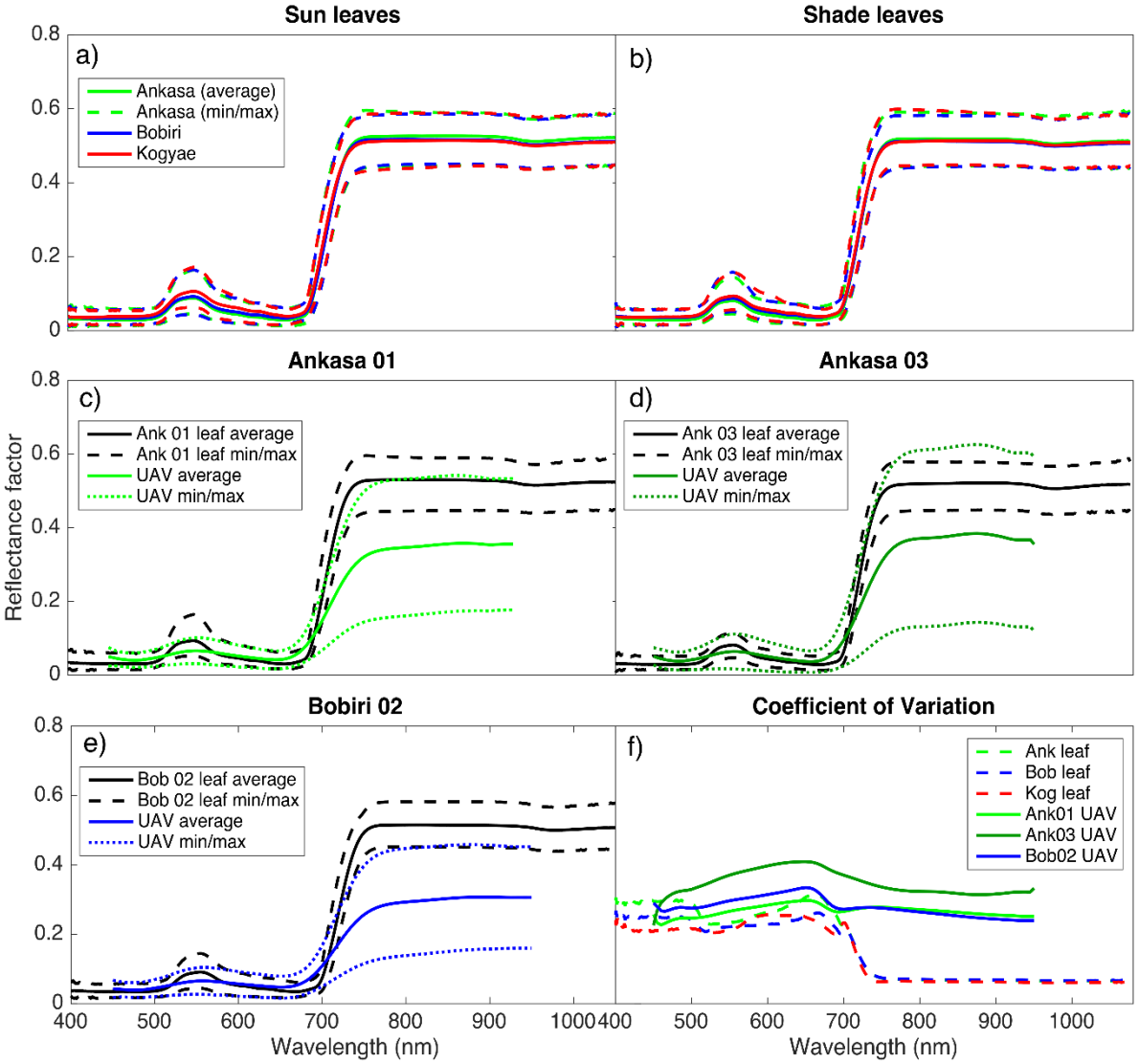


Figure 6.3. (a-b) Mean, maximum and minimum leaf reflectance spectra of all field-collected sun and shade leaves for Ankasa (green), Bobiri (blue) and Kogyae (red). (c–e) Mean, maximum and minimum UAV-collected spectra for Ankasa 01 (bright green), Ankasa 03 (dark green) and Bobiri 02 (blue). The corresponding mean, maximum and minimum field-collected sun leaf spectra for the plot is shown in black for comparison purposes. (f) Mean coefficient of variation of field-collected sun and shade leaves for Ankasa, Bobiri and Kogyae represented by dashed lines. Mean coefficient of variation of the UAV-collected spectra for Ankasa 01, Ankasa 03 and Bobiri 02 represented by solid lines. The colours follow the same colour key above.

PLSR Analyses

In the individual sun leaf PLSR analyses, all leaf traits demonstrated a significant (p < 0.01) relationship with the leaf reflectance spectra (**Table 6.2**). When tested on the independent dataset, LMA and N demonstrated moderate correlations with the leaf spectra (r^2 = 0.42 and 0.43 respectively) and relatively high prediction accuracies of RMSE = 22% and 24%. P and K demonstrated weak but significant correlations with the leaf spectra (r^2 = 0.21 and 0.20) and moderate prediction accuracies of RMSE = 34% and 43% respectively. Ca and Mg displayed both weak correlations (r^2 = 0.23 and 0.14) and low predictive accuracies (RMSE = 60% and 45%).

Shade leaves displayed much stronger spectra–trait correlations than sun leaves, with all traits showing an increase in r^2-val values (**Table 6.2**). In some cases these increases were marked, for example LMA r^2-val increased from 0.42 to 0.55 and P increased from 0.21 to 0.33. RMSE values decreased slightly for LMA and N and otherwise increased or stayed the same for shade leaves.

The bandwidths most significant for predicting foliar traits differed between sun and shade leaves. For sun leaves, 500–700 nm was the most important region for predicting LMA and N. For shade leaves, the NIR was the most important region for predicting LMA and N, with a strong signal at 700 nm for N (**Figure 6.4**. See **Figure S6.1** for shade leaf graphs). Spectral weightings for P mirrored N. Spectral weightings for K, Ca and Mg showed a combination of weightings across the visible, red-edge and NIR regions, indicating that correlations with both LMA and N contributed to their estimates.

When the PLSR analyses were carried out using UAV-collected tree crown spectra, all spectra-trait relationships decreased in strength compared to the individual leaf analyses, with the exception of P and Ca. (**Table 6.2**). For LMA, N, K and Mg, the tree crown spectra were able to predict 5–21% less of the variance between average tree crown trait values than at the individual leaf level. Overall, the tree crown spectra were able to predict LMA, N, P and Ca at p < 0.10. Using UAV-collected spectra, the weightings showed that the NIR region was the most important part of the spectrum for predicting every trait (**Figure 6.4**).

Leaf Type	Sun					Shade					UAV-Collected Spectra				
Spectral Range and Resolution	**Individual Leaves (400–1050 nm) 1 nm FWHM**					**Individual Leaves (400–1050 nm) 1 nm FWHM**					**Tree Crowns (450–950 nm) 9 nm FWHM**				
	Mean ± 1 Std	**RMSE**	**%RMSE**	**r^2-cal**	**r^2-val**	**Mean ± 1 Std**	**RMSE**	**%RMSE**	**r^2-cal**	**r^2-val**	**Mean ± 1 Std**	**RMSE**	**%RMSE**	**r^2-cal**	**r^2-val**
LMA (g m^{-2})	102 ± 35	21.13	22	0.60 ***	0.42 ***	89 ± 33	19.44	20	0.70 ***	0.55 ***	98 ± 20	17.43	18	0.45 ***	0.25 *
N (%)	2.05 ± 0.63	0.48	24	0.78 ***	0.43 ***	2.07 ± 0.64	0.47	23	0.83 ***	0.48 ***	2.15 ± 0.65	0.54	25	0.38 ***	0.22 *
P (%)	0.12 ± 0.05	0.04	34	0.64 ***	0.21 ***	0.12 ± 0.05	0.05	38	0.83 ***	0.33 ***	0.10 ± 0.04	0.03	32	0.38 ***	0.22 *
K (%)	0.89 ± 0.49	0.41	43	0.43 ***	0.20 ***	1.03 ± 0.55	0.48	51	0.52 ***	0.20 ***	0.91 ± 0.44	0.41	44	0.25 *	0.15
Ca (%)	1.39 ± 1.98	0.89	60	0.50 ***	0.23 ***	1.57 ± 1.11	0.94	63	0.66 ***	0.28 ***	0.87 ± 0.66	0.56	64	0.43 ***	0.25 *
Mg (%)	0.34 ± 0.17	0.16	45	0.37 ***	0.14 ***	0.38 ± 0.19	0.18	48	0.45 ***	0.16 ***	0.31 ± 0.18	0.18	57	0.19	0.09

Table 6.2. PLSR results for (i, ii) individual sun and shade leaf spectra. n = 753 for sun and shade leaves respectively. (iii) UAV-collected tree crown spectra. n = 58 for UAV-collected tree crown spectra. RMSE = Root mean square error. r^2-cal represents the 70% of data used to create the empirical model and r^2-val the 30% of data upon which the model was tested. The level of significance (p) is indicated (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Results are the average of 30 model runs. All field-collected data is reported on a dry mass basis.

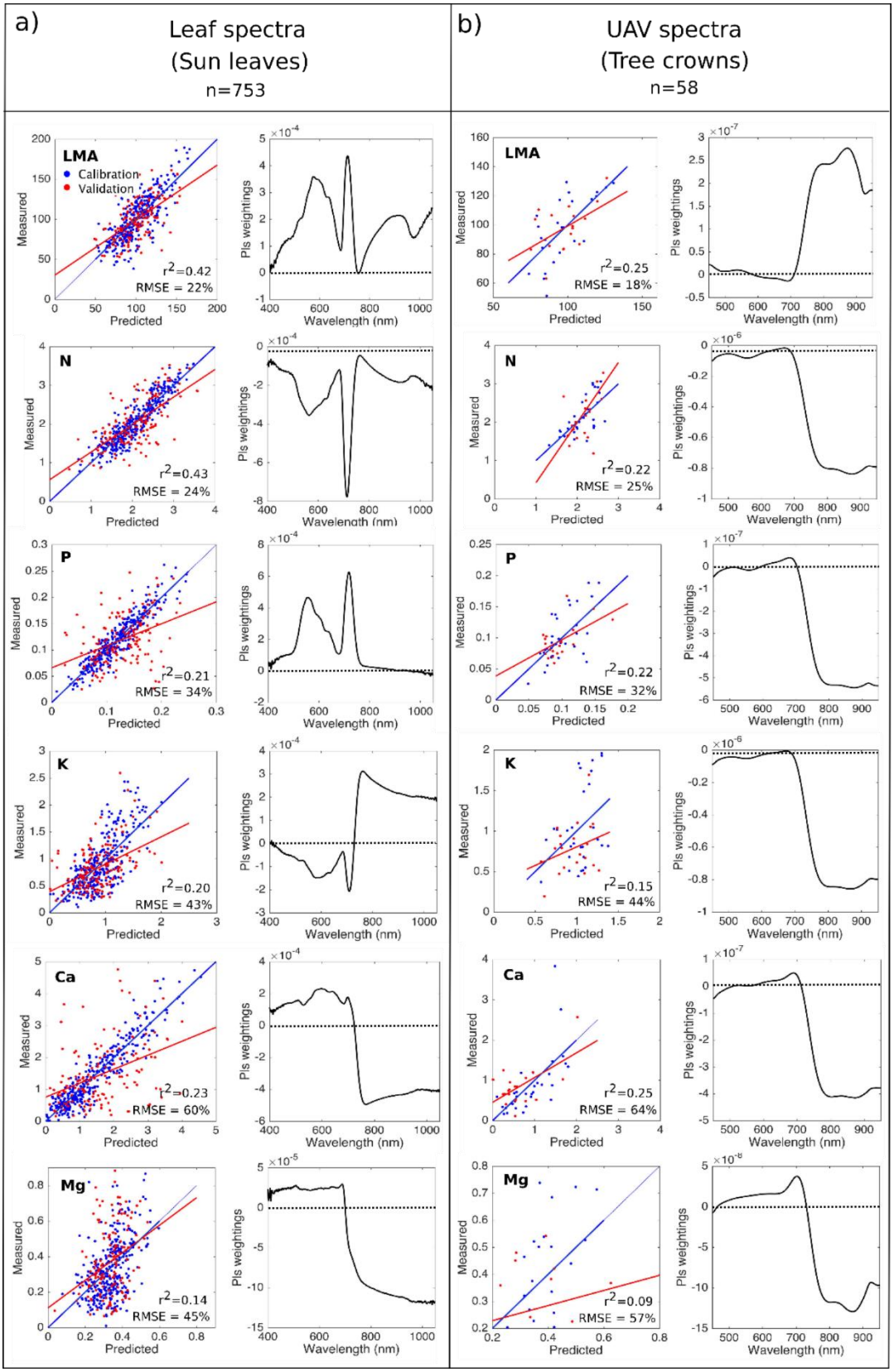


Figure 6.4. (a, left) Relationships between measured and predicted leaf traits using individual sun leaf reflectance spectra and PLS regression analysis. *n* = 753. Blue line = r^2-cal regression line. Red line = r^2-val regression line. RMSE and r^2-val values displayed on graph. (a, right) PLSR spectral weightings for each leaf trait. Deviation from 0 indicates the parts of the spectrum that most strongly influence the empirical model. (b) The same as (**a**), but using UAV-collected reflectance spectra of 58 tree crowns.

Mapping Foliar Traits Using UAV-Collected Hyperspectral Data

The maps of predicted leaf traits (**Figure 6.5**) display similar trait values for Ankasa 01 and Ankasa 03. In Bobiri 02, the maps display much higher N, P and Ca values than in Ankasa. All results were supported by the field-collected values (**Table 6.3**), the one exception being lower LMA in Ankasa 01, which was not found in the weighted field samples.

Similarly to Asner *et al.* (2016c), this study finds that, in general, the method for weighting the field data (basal-area-weighted, stem-abundance-weighted or unweighted mean) had very little effect on the strength, magnitude or direction of the relationships with the UAV data (**Table 6.3**). One exception was LMA, where an unweighted mean resulted in an r^2 value of 1, compared to 0.19 and 0.11 for the basal-area-weighted and stem-abundance-weighted mean respectively. Aside from LMA, however, all field-based trait values showed positive relationships with the remotely sensed plot predictions, with r^2 values ranging from 0.84–1.

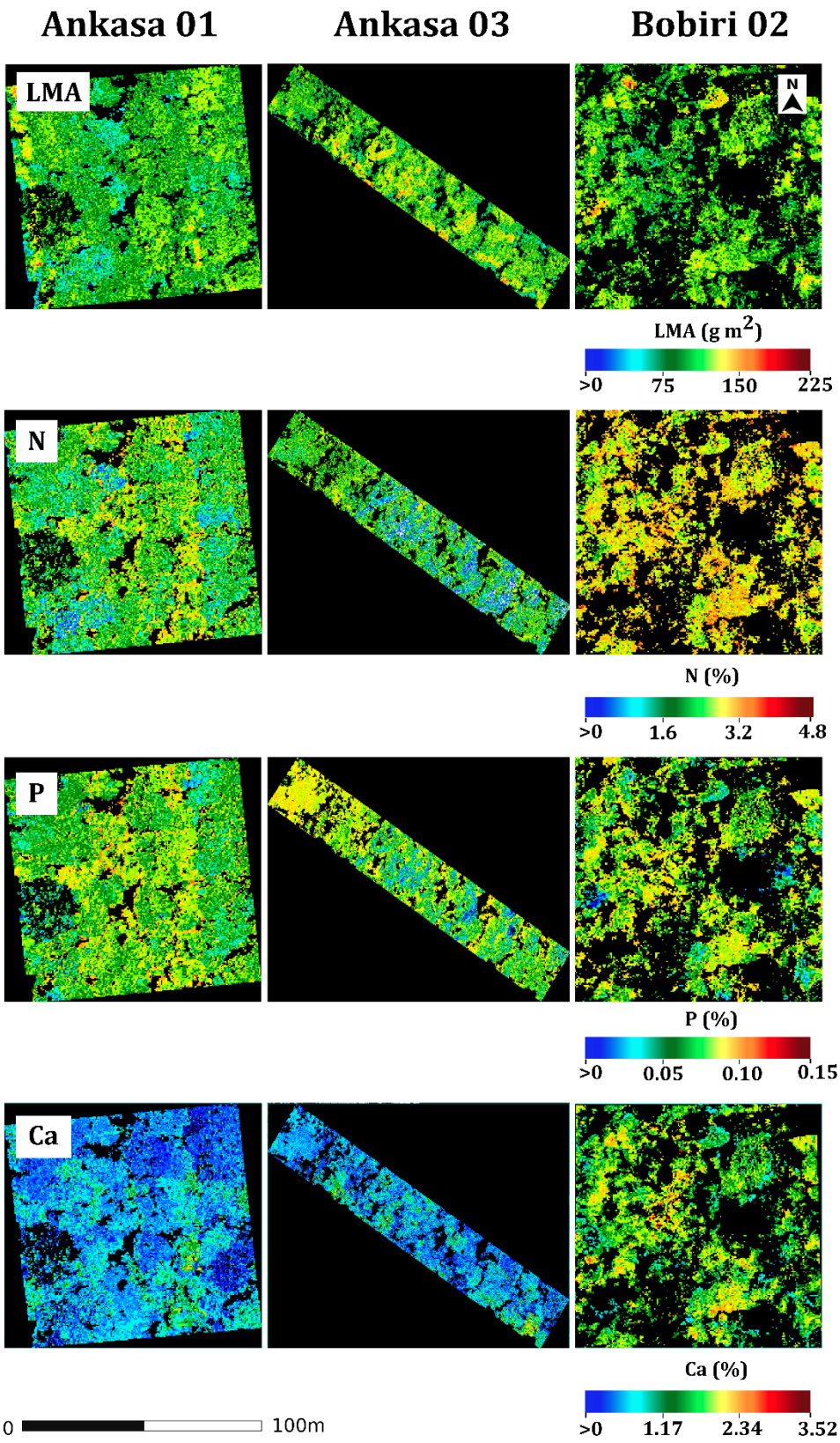


Figure 6.5. Maps of forest canopy LMA, N, P and Ca for three one-ha tropical forest plots in Ghana, West Africa. Created using the regression coefficients generated from the PLSR analyses in Table 2, applied to each pixel of the hyperspectral canopy data using Formula (1). Black areas represent shadows or areas with an NDVI < 0.6. Pixel resolution = 0.5 m.

Leaf Economic Spectrum Trait Interactions

The RGB composite of predicted LMA, N and P showed that the majority of trees in Ankasa 01 and Ankasa 03 fall on the 'slow growing but enduring' end of the LES (**Figure 6.6**). The general hue of both plots is dark blue to purple, representing high LMA, low N and varying P. In contrast, the majority of trees in Bobiri fall on the 'live fast die young' end of the LES, with the plot displaying a general yellow hue, representing low LMA, high N and high P. Within this generality, however, there are still some diverse trait strategies, with prominent tree crowns in Ankasa displaying low LMA, low N and high P, and clusters of foliage in Bobiri displaying high LMA, high N and low P.

At the individual pixel level, no plot displayed a significant relationship between LMA and N values (**Figure 6.7**). All plots displayed a significantly positive correlation ($p < 0.01$) between N and P values ($r^2 = 0.63, 0.59, 0.44$), and a significantly negative correlation between LMA and P values, with Bobiri 02 displaying a comparatively strong relationship of $r^2 = 0.26$. All plots displayed significantly positive correlations between LMA and Ca, and inverse relationships between N and Ca, and P and Ca.

Predicted Trait (Bold) vs. Field-Collected Value	Plot			Statistics		
	Ankasa 01	Ankasa 03	Bobiri 02	RMSE	%RMSE	r^2
Field-sampled basal area (%)	63	56	61			
LMA (g m^{-2})	**90 ± 19** [a]	**100 ± 24** [b]	**98 ± 20** [c]	-	-	-
Basal-area-weighted	102	101	94	7	8	0.19
Stem-abundance-weighted	101	101	92	7	8	0.11
Unweighted	97	101	100	6	5	1 ***
N (%)	**1.97 ± 0.64** [a]	**1.74 ± 0.76** [b]	**2.93 ± 0.50** [c]	-	-	-
Basal-area-weighted	2.06	1.93	2.74	0.16	7	1 ***
Stem-abundance-weighted	2.00	1.89	2.81	0.11	11	1 ***
Unweighted	1.91	2.03	2.45	0.33	15	0.85
P (%)	**0.07 ± 0.02** [a]	**0.07 ± 0.02** [b]	**0.08 ± 0.02** [c]	-	-	-
Basal-area-weighted	0.09	0.10	0.15	0.05	74	0.98
Stem-abundance-weighted	0.09	0.10	0.15	0.05	74	0.98
Unweighted	0.08	0.10	0.13	0.04	62	0.84
Ca (%)	**1.20 ± 0.36** [a]	**1.34 ± 0.48** [b]	**1.87 ± 0.50** [c]	-	-	-
Basal-area-weighted	0.72	0.74	1.82	0.45	30	0.97
Stem-abundance-weighted	0.73	0.73	1.88	0.45	30	0.96
Unweighted	0.71	0.86	1.54	0.44	44	1 ***

Table 6.3. Average predicted foliar traits (bold) for three one-ha tropical forest plots. Calculated using UAV-collected hyperspectral data and PLSR analysis. Field-collected trait values are shown for comparison purposes. The field-collected data have been weighted using three methods: basal-area-weighting, stem-abundance-weighting and no weighting. The root mean square error (RMSE), percentage root mean square error (%RMSE) and r^2 values of predicted vs. field-collected traits are shown for each method of weighting. (n = 3). The level of significance (p) is indicated (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Significant differences ($p < 0.01$) between the predicted trait values of each plot are denoted using the letters a–c Where necessary, values have been rounded to two decimal places.

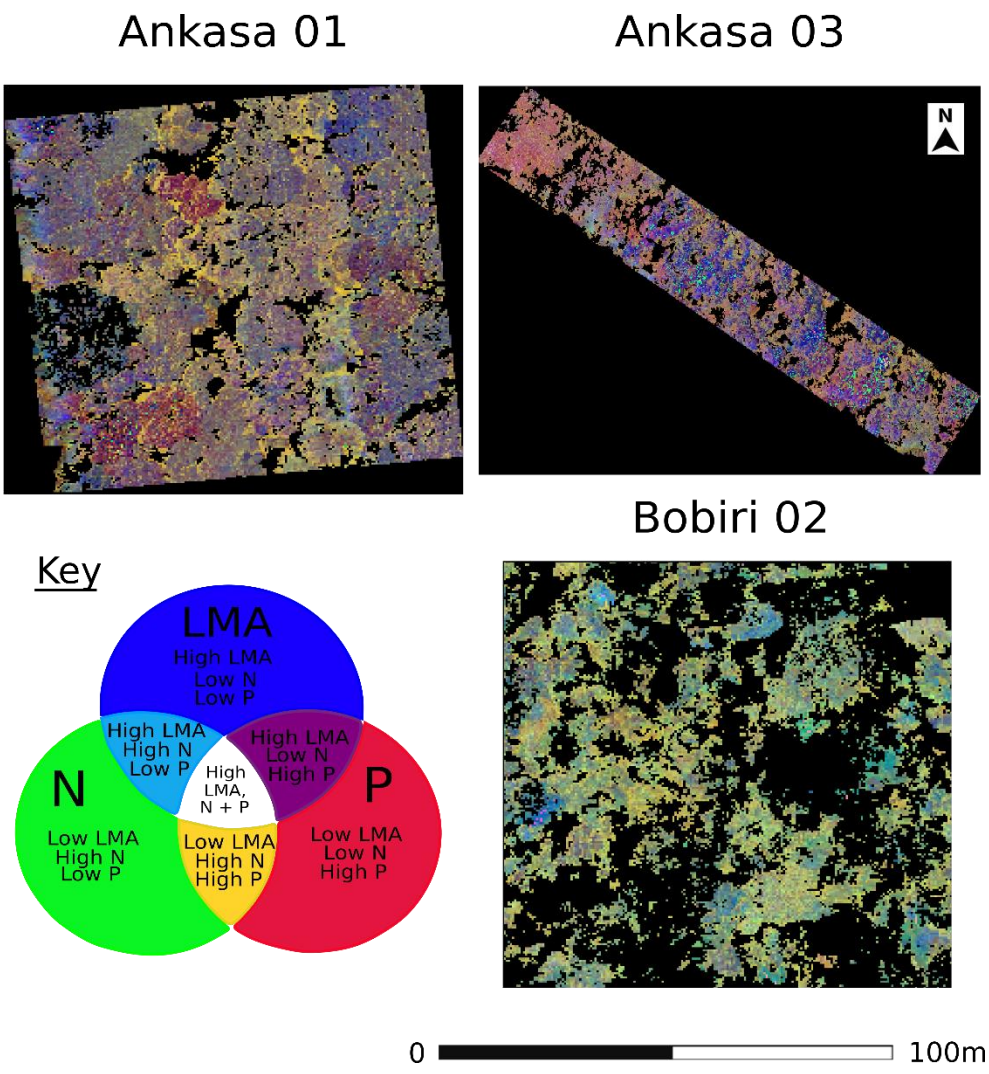


Figure 6.6. RGB composite of LES trait interactions (red = P, green = N, blue = LMA). Produced using the maps of predicted canopy LMA, N and P in **Figure 6.4**. Black areas represent shadows or areas with an NDVI < 0.6. In the Venn diagram, the continuous rgb colour scheme is shown as discrete categories for ease of interpretation. Pixel resolution = 0.5 m.

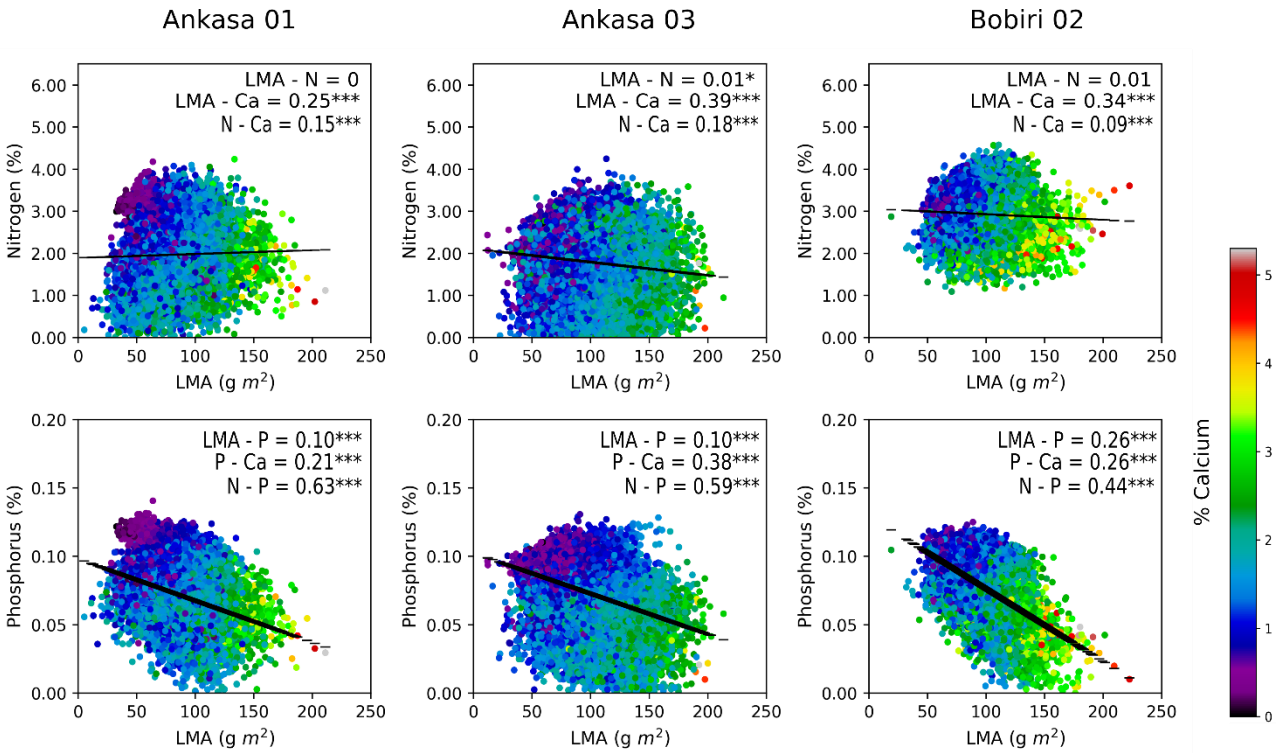


Figure 6.7. Relationship between LMA–N and LMA–P for each pixel of **Figure 6.6**. Colour as a function of Ca. r^2 values are shown on the graphs with the level of significance (p) indicated (* $p < 0.10$; ** $p < 0.05$; ** $p < 0.01$). n = 34789 for Ankasa 01. n = 32807 for Ankasa 03. n = 19014 for Bobiri 02. These graphs are based on the remotely sensed values, not field-collected data.

6.7. Discussion

This study demonstrates the potential of using leaf spectral properties and UAVs to scale up from individual leaves to landscapes. Calibrated on a field campaign that measured the trait values of 753 sun leaves, we used UAV-collected VNIR hyperspectral data to map LMA, N, P and Ca across three hectares of diverse tropical forest, revealing patterns of trait variation both between and within forest sites. While this study limited itself to analysing interactions between LES traits and validating the methodology, it demonstrates the potential of UAV-collected spectra for scaling up similar leaf trait-based work. For example, LMA and leaf chemical traits have been shown to be good indicators of canopy phylogeny (Asner *et al.* 2014a; Asner and Martin, 2011), invasive species (Asner *et al.* 2008), decomposition rates (Bakker *et al.* 2011; Hättenschwiler *et al.* 2011), soil fertility (Ordoñez *et al.* 2009) and drought tolerance (Poorter and Markesteijn, 2008).

When compared to the individual leaf spectra, the UAV-collected spectra displayed a similar shape, but lower reflectance in the visible and NIR regions (**Figure 6.3**). Increased leaf area index and variable leaf angle distribution in the canopy causes multiple scattering and the increased absorption of light compared to single leaves. The UAV-collected spectra also displayed significantly higher variation than the hand-collected spectra. Higher variation in the UAV spectra is almost certainly due to the far higher number of spectra pixels recorded across each plot—34,789 pixels across Ankasa 01, compared to 352 hand-collected leaf spectra. While only healthy, mature sun and shade leaves were measured by hand, the UAV captured all visible sun foliage across the plot, including leaves of different ages, health and species. Leaf age has been shown to have a significant effect on spectral properties (Chavana-Bryant *et al*. 2017), as has leaf health (Gitelson *et al.* 2003; West *et al.* 2003). The individual sun and shade leaf spectra were collected from 60 species representing 27 families across the three plots. While these species represented >55% of the basal area of each plot, there were a total of 130 species from 51 families present across the three plots, contributing to far higher spectral variation in the UAV spectral data.

As expected from our first hypothesis, significant relationships were found between LMA, N, P, K, Ca, Mg and leaf reflectance spectra at the individual leaf level (**Table 6.2**). These

were strongest for the directly estimated traits of LMA and N. The covariate relationships found between LMA, N and the indirectly expressed nutrients are discussed in more detail below. Results were slightly lower than those found by Asner *et al.* (2011b), who used a very similar spectral range and method, but a much larger and more diverse sample size, suggesting that high ecological variety drives stronger PLSR relationships. Thus, our first hypothesis *is accepted*.

Across all traits, shade leaves demonstrated much stronger spectral relationships than sun leaves (**Table 6.2**). In the spectroscopy literature, shade leaves are chronically understudied. Among the few studies that include shade leaves in their analysis however, this finding does not appear to be unique. Using PLSR to predict leaf age in a Peruvian and Brazilian forest, Wu *et al.* (2017) found that on a species-by-species basis, shade leaves predicted 5–27% more of the variance in LMA than sun leaves and 6–28% more of the variance in leaf water content.

The higher performance of shade leaves may be linked to their structure. To increase light capture efficiency, shade leaves are thinner, possess fewer cells per leaf area than sun leaves and have palisade cells of roughly equal dimensions that concentrate chloroplasts in a dense layer on the leaf's adaxial surface. In contrast, sun leaves are thicker, with multiple mesophyll layers and columnar palisade cells to allow light to pass through vacuoles and the spaces between cells and reach chloroplasts in the spongy mesophyll (Lee *et al.* 1990; Ellsworth and Reich, 1993). Thus, a beam of light penetrating a sun leaf has more cell–wall–air interfaces to pass through and a longer path length, increasing the probability of being scattered and absorbed, and potentially decreasing the signal-to-noise ratio of the reflectance spectra. The fact that shade leaf–trait relationships are equal or stronger than sun leaf–trait relationships is important, because shade leaves constitute the majority of tropical foliage and influence canopy spectral signatures in the NIR—hence, their spectral contribution will assume greater importance when the spectral range of the UAV-collected data is extended up to 2500 nm.

Scaling up to the airborne level, LMA, N, P and Ca demonstrated a significant (at $p < 0.10$) relationship with the UAV-collected tree crown spectra (**Table 6.2**). These relationships were generally lower than those at found at the individual leaf level. Thus, our second

hypothesis is *rejected.* Lower relationships may be due to several reasons. Firstly, the PLSR analyses were carried out with a much smaller sample size of tree crowns—58 tree crowns vs. 753 individual leaves. Although PLSR analyses have been carried out with similar sample sizes (e.g. Doughty *et al.* 2011), results have been shown to improve with increasing sample size, until stabilization is reached at n = 1000 (Asner *et al.* 2011b). Secondly, the airborne canopy data were collected 52–64 weeks after the field collection campaign. Interannual variability of leaf traits has been found to be low for LMA and leaf nutrient values collected at the same time of year or in the same season (Garnei *et al.* 2001; Riipi *et al.* 2004; Gotsch *et al.* 2010). Yet, when the airborne canopy data were collected, tree crown trait values may still have altered slightly from their field-collected values. As the PLSR method works by calibrating and predicting values based on spectral variability, slight inaccuracies in crown trait values would weaken the PLSR results. It is important to note that we do not list these factors to undermine the study, but rather to give further confidence in its results. The fact that, notwithstanding these limitations, we found significant spectra–tree crown relationships for 4 out of 6 traits, supporting the value of using UAVs to estimate canopy traits.

In addition to the factors above, we hypothesize that the largest improvements in airborne spectra–trait relationships will be made when the SWIR (700–2500 nm) region is included in the canopy spectra measurements. Using the full VSWIR range has been shown to significantly increase spectra–trait relationships (Asner *et al*. 2011b; Ebbers *et al.* 2002; Kokaly *et al.* 2009; Skidmore *et al.* 2010), as the majority of traits have significant expressions in the short-wave infrared spectrum. Airborne hyperspectral sensors covering the full 400–2500 nm range are commercially available (Blaaberg *et al*. 2014). Alternatively, combining information from multiple hyperspectral sensors is possible (Asner *et al.* 2012b), and would also provide all the relevant bands. Experimenting with the expanding field of nonlinear transformation methods (often referred to as machine learning algorithms) may also improve retrieval rates (Verrelst *et al.* 2018), although in this study PLSR was used as it offers a more transparent, physical approach to quantifying spectra-trait relationships and facilitates comparisons with similar studies.

At the whole plot level, the UAV-predicted plot averages showed good agreement with field-collected values (**Table 6.3**). With the exception of LMA, basal area- and stem-

abundance weighting of field-collected traits generally demonstrated similar or higher r^2 values, and similar or lower %RMSE values, than the unweighted means, supporting the use of weighting field data when calibrating or validating remotely sensed variables. Despite the fact that few regression models were statistically significant, the lack of statistical significance is more likely due to the extremely small sample size (3 data points) than to a lack of agreement between the field-collected and remotely sensed values. Weighted LMA was the only field-collected trait not to demonstrate a positive relationship with the remotely sensed canopy values. However, unlike N, P or Ca, variations in LMA were small between the three sites. The remotely sensed values demonstrated an uncertainty of 19–24 g m^2—larger than the variation observed between the three sites. Thus, it is unsurprising that the field and remote sensing approaches do not agree on a mean LMA trend for these 3 plots—greater LMA variation is required.

The fact that the UAV-collected data were able to quantify average trait values for each plot represents an important step in scaling up these results. One-ha tropical forest study plots are not always representative of their host landscapes (Asner *et al.* 2016c; Marvin *et al.* 2014). Thus, while UAVs could be calibrated on field results collected from smaller study sites, their greatest asset in is mapping these values over larger areas, where field-collection is impractical.

Across the three sites, the observed three-way interaction between LMA, N and P supports the proposal of a multi-trait LES, sensitive to climatic and edaphic constraints (Asner *et al.* 2016b; Hidaka *et al.* 2011; Asner *et al.* 2014b). From the RGB composite, it can be seen that Ankasa 01 and Ankasa 03 generally display high LMA, low N and low P values, while Bobiri 02 generally exhibits low LMA, high N and high P values. These trait patterns are almost certainly due to the strong rainfall gradient found across West Africa. While increasing precipitation is generally associated with more favourable conditions and trait trade-offs that favour fast growth and short mortality, at high precipitation rates (over 2000 mm), nutrient gradients have been found to interact with precipitation gradients. High precipitation causes soil nutrient leaching, promoting the selection of species for traits on the persistent end of the economics spectrum (Santiago *et al*. 2004a, 2004b; Maharjan *et al.* 2011). Evergreenness has been found to be a response to the relatively low nutrient availability in very wet tropical forest, as increased leaf longevity

is an effective nutrient conservation mechanism, resulting in a longer photosynthetic season and a reduced demand on the amount of soil nutrients required to replace leaves (Antonio Vazquez and Givnish, 1998; Cunningham *et al.* 1999; Malhi *et al.* 2015). Hence, with their weathered tropical soils promoting the growth of evergreen species, Ankasa 01 and Ankasa 03 demonstrate high LMA values, high leaf longevity (9–12 months) and lower N, P and Ca values. These leaf strategies are not limited to plant foliage, but are mirrored by the tree's overall performance. According to the carbon cycling data for the plots, Ankasa demonstrates high wood residence rates (99 ± 22.1 years in Ankasa 01) and lower NPP (13.3 ± 0.6 Mg C ha^{-1} $year^{-1}$: all data from Moore *et al.* (2017) **Table 6.1**).

In contrast, the RGB composite shows that the most common trait strategy in Bobiri represents the 'live fast die young' end of the LES spectrum. Here, lower mean annual rainfall reduces soil leaching rates, while still providing enough water for fast growth. Photosynthetic rates in Bobiri 02 are correspondingly higher, with an average A_{sat} rate of 7.92 μmol m^{-2} s^{-1}, compared to an average of 5.65 for Ankasa 01 (Gvozdevaite *et al.* 2018). Richer soils support a higher number of deciduous tree species—if the resources to build new organic matter are there, there are no constraints to adopting higher rates of leaf production throughout the year. Thus, leaf longevity in Bobiri is lower than in Ankasa (5–6 months), along with wood residence rates (32 ± 7.16 years), while NPP is higher (16.7 ± 0.6 Mg C ha^{-1} $year^{-1}$: all data from Moore *et al.* (2017). See **Table 6.1**).

In contrast to the majority of leaf trait studies, our findings were unique in that we did not find a significant relationship between LMA and N across any of the three plots. Here, these relationships are based on the remotely sensed data, yet the same has been found in the field-data (Gvozdevaite *et al.* 2018). Thus, our third hypothesis is *rejected*. This finding is probably due to the small spatial distribution of our data. While most studies analysing LMA–N interactions collect data over environmental gradients (Meir *et al.* 2001; van der Weg *et al.* 2009; Li *et al.* 2015; de la Riva *et al.* 2016), our study focused on trait interactions within fairly homogenous one-ha plots. Thus, while LMA and N values may respond to similar environmental variables over wider areas, under more constant environmental conditions, they appear to be ordered by different controls and do not correlate at the individual leaf level.

Significant relationships were found between all other traits. Bobiri 02 displayed a comparatively strong relationship between LMA and P (r^2 = 0.26), supporting research that finds, when soil P is high, phylogeny exerts a stronger control on foliar P than when soil P is limited (Asner *et al*. 2014, 2014). As both Ankasa sites also display relatively high soil P values compared to many Amazonian soils, a phylogentic ordering of P along the same lines as a phylogenetic ordering of LMA may explain why significant relationships were also found between LMA and P at these sites. As widely reported in the LES literature, N and P exhibited strongly positive relationships in each plot, lending weight to the idea of a stoichiometric link between N and P (Hidaka *et al.* 2011; van der Weg *et al.* 2009; Porder *et al.* 2005). LMA and Ca also demonstrated significantly positive relationships in each plot, probably due to the role of Ca in foliar cell wall development (Demarty *et al.* 1984)—as LMA increases, more Ca is required.

Overall, our results support the use of a multi-trait LES incorporating P and Ca values (Wright *et al*. 2004; Asner *et al*. 2017). In accordance with Asner *et al*. (2016b), we find that environmental filters generate more diverse plant trait strategies than be captured by a universal leaf LMA–N relationship and caution against using simple N–LMA tradeoffs to analyse ecosystem function.

6.8. Conclusions

In this study, we show that LMA and multiple leaf chemical traits can be retrieved from individual leaf and plot level spectra using hand-held instruments and UAVs. Retrieval of foliar traits at the tree crown level was weaker. The next steps are to mount both a VNIR and SWIR sensor onto a UAV to increase trait estimation accuracy and attempt to expand the range of quantifiable traits. The next generation of hyperspectral satellites—EnMAP, FLEX, HISUI, HyspIRI, PRISMA, SHALOM (Stuffler *et al.* 2007; Moreno *et al.* 2006; Iwasaki *et al.* 2011; Lee *et al.* 2015; Pignatti *et al.* 2013; Natale *et al.* 2013)—will be critical in providing global coverage and monitoring of functional diversity. In the meantime, developing cheaper and easier ways to map leaf traits in tropical forests will provide invaluable data for understanding LES traits and their interactions, as well as helping to inform the next generation of dynamic vegetation and climate models (Sakschewski *et al.* 2015; Scheiter *et al.* 2013).

6.9. Acknowledgements

This work is a product of the Global Ecosystems Monitoring (GEM) network (**gem.tropicalforests.ox.ac.uk**). The UAV data was collected by Wageningen UR Unmanned Aerial Remote Sensing Facility (**http://www.wageningenur.nl/uarsf**). We thank the FORIG and Ankasa National Park for providing logistical support.

6.10. Funding

This research was funded by a Google Earth Engine Research Award to C.E.D., E.R.T. and A.J.A. (PRQ22909209), and a NERC DTP student grant to E.R.T. (NE/L002612/1). The UAV data collection was funded by INTERREG V A-project SPECTORS (Project number: 143081). The HYMSY development was funded by INTERREG project 'Smart Inspectors' (Project number: I-1-03 = 176). The field-collected trait data was funded by the Royal Society–Leverhulme Africa Capacity Building Award (AA130026), and an Advanced Investigator Award to Y.M. from the European Research Council (GEM-TRAIT; Grant No. 321131). Y.M. is supported by the Jackson Foundation. The APC was funded by an Oxford RCUK Open Access Block Grant.

6.11. References

Antonio Vazquez G., J. and Givnish, T.J. (1998) Altitudinal gradients in tropical forest composition, structure, and diversity in the Sierra de Manantlán. *Journal of Ecology,* , pp. 999-1020.

Asner, G.P., Jones, M.O., Martin, R.E., Knapp, D.E. and Hughes, R.F. (2008) Remote sensing of native and invasive species in Hawaiian forests. *Remote Sensing of Environment,* 112 (5), pp. 1912-1926.

Asner, G.P., Martin, R.E., Ford, A.J., Metcalfe, D.J. and Liddell, M.J. (2009) Leaf chemical and spectral diversity in Australian tropical forests. *Ecological Applications,* 19 (1), pp. 236-253.

Asner, G.P., Martin, R.E., Tupayachi, R., Emerson, R., Martinez, P., Sinca, F., Powell, G.V.N., Wright, S.J. and Lugo, A.E. (2011a) Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. *Ecological Applications,* 21 (1), pp. 85-98.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C., Carranza, L., Martinez, P., Houcheime, M., Sinca, F. and Weiss, P. (2011b) Spectroscopy of canopy chemicals in humid tropical forests. *Remote Sensing of Environment,* 115 (12), pp. 3587-3598.

Asner, G.P., Martin, R.E. and Suhaili, A.B. (2012a) Sources of canopy chemical and spectral diversity in lowland Bornean forest. *Ecosystems,* 15 (3), pp. 504-517.

Asner, G.P., Knapp, D.E., Boardman, J., Green, R.O., Kennedy-Bowdoin, T., Eastwood, M., Martin, R.E., Anderson, C. and Field, C.B. (2012b) Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. *Remote Sensing of Environment,* 124, pp. 454-465.

Asner, G.P., Martin, R.E., Tupayachi, R., Anderson, C.B., Sinca, F., Carranza-Jiménez, L. and Martinez, P. (2014a) Amazonian functional diversity from forest canopy chemical assembly. *Proceedings of the National Academy of Sciences,* 111 (15), pp. 5604-5609.

Asner, G.P., Martin, R.E., Carranza-Jiménez, L., Sinca, F., Tupayachi, R., Anderson, C.B. and Martinez, P. (2014b) Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to Amazon region. *New Phytologist,* 204 (1), pp. 127-139.

Asner, G.P., Martin, R.E., Anderson, C.B. and Knapp, D.E. (2015) Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sensing of Environment,* 158, pp. 15-27.

Asner, G.P. and Martin, R.E. (2016a) Spectranomics: Emerging science and conservation opportunities at the interface of biodiversity and remote sensing. *Global Ecology and Conservation,* 8, pp. 212-219.

Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E. and Vaughn, N. (2016b) Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proceedings of the National Academy of Sciences,* 113 (28), pp. E4043-E4051.

Asner, G.P., Martin, R.E., Anderson, C.B., Kryston, K., Vaughn, N., Knapp, D.E., Bentley, L.P., Shenkin, A., Salinas, N. and Sinca, F. (2016c) Scale dependence of canopy trait distributions along a tropical forest elevation gradient. *New Phytologist,* 214 (3), pp. 973-988.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn, N.R. and Llactayo, W. (2017) Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science,* 355 (6323), pp. 385-389.

Asner, G.P. and Martin, R.E. (2011) Canopy phylogenetic, chemical and spectral assembly in a lowland Amazonian forest. *New Phytologist,* 189 (4), pp. 999-1012.
Bakker, M.A., Carreño-Rocabado, G. and Poorter, L. (2011) Leaf economics traits predict litter decomposition of tropical plants and differ among land use types. *Functional Ecology,* 25 (3), pp. 473-483.
Blaaberg, S., Løke, T., Baarstad, I., Fridman, A. and Koirala, P. (2014) HySpex ODIN-1024: a new high-resolution airborne HSI system. International Society for Optics and Photonics, pp. 90700L.
Brienen, R.J.W., Phillips, O.L., Feldpausch, T.R., Gloor, E., Baker, T.R., Lloyd, J., Lopez-Gonzalez, G., Monteagudo-Mendoza, A., Malhi, Y. and Lewis, S.L. (2015) Long-term decline of the Amazon carbon sink. *Nature,* 519 (7543), pp. 344-348.
Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G.P., Anastasiou, A., Enquist, B.J., Cosio Caravasi, E.,G., Doughty, C.E., Saleska, S.R. and Martin, R.E. (2017) Leaf aging of Amazonian canopy trees as revealed by spectral and physiochemical measurements. *New Phytologist,* 214 (3), pp. 1049-1063.
Cho, M.A., Mathieu, R., Asner, G.P., Naidoo, L., Van Aardt, J., Ramoelo, A., Debba, P., Wessels, K., Main, R. and Smit, I.P.J. (2012) Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment,* 125, pp. 214-226.
Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J. and Heimann, M. (2014) Carbon and other biogeochemical cycles. In *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 University Press, pp. 465-570.
Clark, M.L., Roberts, D.A. and Clark, D.B. (2005) Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment,* 96 (3-4), pp. 375-398.
Cleveland, C.C., Houlton, B.Z., Smith, W.K., Marklein, A.R., Reed, S.C., Parton, W., Del Grosso, S.,J. and Running, S.W. (2013) Patterns of new versus recycled primary production in the terrestrial biosphere. *Proceedings of the National Academy of Sciences,* 110 (31), pp. 12733-12737.
Cleveland, C.C., Taylor, P., Chadwick, K.D., Dahlin, K., Doughty, C.E., Malhi, Y., Smith, W.K., Sullivan, B.W., Wieder, W.R. and Townsend, A.R. (2015) A comparison of plot-based satellite and Earth system model estimates of tropical forest net primary production. *Global Biogeochemical Cycles,* 29 (5), pp. 626-644.
Cornwell, W.K., Cornelissen, J.H.C., Amatangelo, K., Dorrepaal, E., Eviner, V.T., Godoy, O., Hobbie, S.E., Hoorens, B., Kurokawa, H. and Pérez-Harguindeguy, N. (2008) Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecology Letters,* 11 (10), pp. 1065-1071.
Cunningham, S.A., Summerhayes, B. and Westoby, M. (1999) Evolutionary divergences in leaf structure and chemistry, comparing rainfall and soil nutrient gradients. *Ecological Monographs,* 69 (4), pp. 569-588.
Curran, P.J. (1989) Remote sensing of foliar chemistry. *Remote Sensing of Environment,* 30 (3), pp. 271-278.
Dalponte, M., Ene, L.T., Marconcini, M., Gobakken, T. and Næsset, E. (2015) Semi-supervised SVM for individual tree crown species classification. *ISPRS Journal of Photogrammetry and Remote Sensing,* 110, pp. 77-87.

de la Riva, E.,G., Olmo, M., Poorter, H., Ubera, J.L. and Villar, R. (2016) Leaf mass per area (LMA) and its relationship with leaf structure and anatomy in 34 Mediterranean woody species along a water availability gradient. *PloS one,* 11 (2), pp. e0148788.
Demarty, M., Morvan, C. and Thellier, M. (1984) Calcium and the cell wall. *Plant, Cell & Environment,* 7 (6), pp. 441-448.
Díaz, S., Hodgson, J.G., Thompson, K., Cabido, M., Cornelissen, J.H.C., Jalili, A., Montserrat-Martí, G., Grime, J.P., Zarrinkamar, F. and Asri, Y. (2004) The plant traits that drive ecosystems: evidence from three continents. *Journal of vegetation science,* 15 (3), pp. 295-304.
Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C. and Prentice, I.C. (2016) The global spectrum of plant form and function. *Nature,* 529 (7585), pp. 167-171.
Dixon, R.K., Solomon, A.M., Brown, S., Houghton, R.A., Trexier, M.C. and Wisniewski, J. (1994) Carbon pools and flux of global forest ecosystems. *Science,* 263 (5144), pp. 185-190.
Doughty, C.E., Asner, G.P. and Martin, R.E. (2011) Predicting tropical plant physiology from leaf and canopy spectroscopy. *Oecologia,* 165 (2), pp. 289-299.
Drenovsky, R.E., Grewell, B.J., D'antonio, C.,M., Funk, J.L., James, J.J., Molinari, N., Parker, I.M. and Richards, C.L. (2012) A functional trait perspective on plant invasion. *Annals of botany,* 110 (1), pp. 141-153.
Ebbers, M., Wallis, I.R., Dury, S., Floyd, R. and Foley, W.J. (2002) Spectrometric prediction of secondary metabolites and nitrogen in fresh Eucalyptus foliage: towards remote sensing of the nutritional quality of foliage for leaf-eating marsupials. *Australian Journal of Botany,* 50 (6), pp. 761-768.
Ellsworth, D.S. and Reich, P.B. (1993) Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. *Oecologia,* 96 (2), pp. 169-178.
Enquist, B.J., Norberg, J., Bonser, S.P., Violle, C., Webb, C.T., Henderson, A., Sloat, L.L. and Savage, V.M. (2015) Scaling from traits to ecosystems: developing a general trait driver theory via integrating trait-based and metabolic scaling theories. *Advances in Ecological Research,* 52, pp. 249-318.
Garnier, E., Laurent, G., Bellmann, A., Debain, S., Berthelier, P., Ducout, B., Roumet, C. and Navas, M. (2001) Consistency of species ranking based on functional leaf traits. *New Phytologist,* 152 (1), pp. 69-83.
Gates, D.M., Keegan, H.J., Schleter, J.C. and Weidner, V.R. (1965) Spectral properties of plants. *Applied Optics,* 4 (1), pp. 11-20.
Gitelson, A.A., Gritz, Y. and Merzlyak, M.N. (2003) Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology,* 160 (3), pp. 271-282.
Gökkaya, K., Thomas, V., Noland, T.L., McCaughey, H., Morrison, I. and Treitz, P. (2015) Prediction of macronutrients at the canopy level using spaceborne imaging spectroscopy and LiDAR data in a mixedwood boreal forest. *Remote Sensing,* 7 (7), pp. 9045-9069.
Gotsch, S.G., Powers, J.S. and Lerdau, M.T. (2010) Leaf traits and water relations of 12 evergreen species in Costa Rican wet and dry forests: patterns of intra-specific variation across forests and seasons. *Plant Ecology,* 211 (1), pp. 133-146.
Grace, J., Malhi, Y., Higuchi, N. and Meir, P. (2001) Productivity of tropical rain forests. *Terrestrial global productivity,* , pp. 401-426.

Gvozdevaite, A., Oliveras, I., Domingues, T.F., Peprah, T., Boakye, M., Afriyie, L., da Silva Peixoto, K., de Farias, J., Almeida de Oliveira, E. and Almeida Farias, C.C. (2018) Leaf-level photosynthetic capacity dynamics in relation to soil and foliar nutrients along forest–savanna boundaries in Ghana and Brazil. *Tree physiology,* 38 (12), pp. 1912-1925.
Haaland, D.M. and Thomas, E.V. (1988) Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry,* 60 (11), pp. 1193-1202.
Hättenschwiler, S., Coq, S., Barantal, S. and Handa, I.T. (2011) Leaf traits and decomposition in tropical rainforests: revisiting some commonly held views and towards a new hypothesis. *New Phytologist,* 189 (4), pp. 950-965.
Hidaka, A. and Kitayama, K. (2011) Allocation of foliar phosphorus fractions and leaf traits of tropical tree species in response to decreased soil phosphorus availability on Mount Kinabalu, Borneo. *Journal of Ecology,* 99 (3), pp. 849-857.
Hoffmann, W.A., Franco, A.C., Moreira, M.Z. and Haridasan, M. (2005) Specific leaf area explains differences in leaf traits between congeneric savanna and forest trees. *Functional Ecology,* 19 (6), pp. 932-940.
Huang, C. and Asner, G.P. (2009) Applications of remote sensing to alien invasive plant studies. *Sensors,* 9 (6), pp. 4869-4889.
Iwasaki, A., Ohgi, N., Tanii, J., Kawashima, T. and Inada, H. (2011) Hyperspectral Imager Suite (HISUI)-Japanese hyper-multi spectral radiometer. IEEE, pp. 1025.
Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F.W., Asner, G.P., Guralnick, R., Kattge, J., Latimer, A.M. and Moorcroft, P. (2016) Monitoring plant functional diversity from space. *Nature plants,* 2 (3), pp. 1-5.
Jobbágy, E.,G. and Jackson, R.B. (2000) The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological Applications,* 10 (2), pp. 423-436.
Kattge, J., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P.B. and Wright, I.J. (2011) TRY–a global database of plant traits. *Global Change Biology,* 17 (9), pp. 2905-2935.
Knipling, E.B. (1970) Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment,* 1 (3), pp. 155-159.
Kokaly, R.F. (2001) Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment,* 75 (2), pp. 153-161.
Kokaly, R.F., Asner, G.P., Ollinger, S.V., Martin, M.E. and Wessman, C.A. (2009) Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment,* 113, pp. S78-S91.
Lavorel, S., Grigulis, K., Lamarque, P., Colace, M., Garden, D., Girel, J., Pellet, G. and Douzet, R. (2011) Using plant functional traits to understand the landscape distribution of multiple ecosystem services. *Journal of Ecology,* 99 (1), pp. 135-147.
Lavorel, S., Storkey, J., Bardgett, R.D., De Bello, F., Berg, M.P., Le Roux, X., Moretti, M., Mulder, C., Pakeman, R.J. and Díaz, S. (2013) A novel framework for linking functional diversity of plants with other trophic levels for the quantification of ecosystem services. *Journal of Vegetation Science,* 24 (5), pp. 942-948.
Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L. and Shreenan, R. (2003) Combined high-density lidar and multispectral imagery for individual tree crown analysis. *Canadian Journal of Remote Sensing,* 29 (5), pp. 633-649.
Lee, C.M., Cable, M.L., Hook, S.J., Green, R.O., Ustin, S.L., Mandl, D.J. and Middleton, E.M. (2015) An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. *Remote Sensing of Environment,* 167, pp. 6-19.

Lee, D.W., Bone, R.A., Tarsis, S.L. and Storch, D. (1990) Correlates of leaf optical properties in tropical forest sun and extreme-shade plants. *American Journal of Botany,* 77 (3), pp. 370-380.
Lewis, S.L., Lopez-Gonzalez, G., Sonké, B., Affum-Baffoe, K., Baker, T.R., Ojo, L.O., Phillips, O.L., Reitsma, J.M., White, L. and Comiskey, J.A. (2009) Increasing carbon storage in intact African tropical forests. *Nature,* 457 (7232), pp. 1003-1006.
Maharjan, S.K., Poorter, L., Holmgren, M., Bongers, F., Wieringa, J.J. and Hawthorne, W.D. (2011) Plant functional traits and the distribution of West African rain forest trees along the rainfall gradient. *Biotropica,* 43 (5), pp. 552-561.
Malhi, Y., Doughty, C.E., Goldsmith, G.R., Metcalfe, D.B., Girardin, C.A.J., Marthews, T.R., del Aguila-Pasquel, J., Aragão, L.,E.O.C., Araujo-Murakami, A. and Brando, P. (2015) The linkages between photosynthesis, productivity, growth and biomass in lowland Amazonian forests. *Global Change Biology,* 21 (6), pp. 2283-2295.
Malhi, Y. and Grace, J. (2000) Tropical forests and atmospheric carbon dioxide. *Trends in Ecology & Evolution,* 15 (8), pp. 332-337.
Marvin, D.C., Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E., Sinca, F. and Tupayachi, R. (2014) Amazonian landscapes and the bias in field studies of forest structure and biomass. *Proceedings of the National Academy of Sciences,* 111 (48), pp. E5224-E5232.
McGill, B.J., Enquist, B.J., Weiher, E. and Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in ecology & evolution,* 21 (4), pp. 178-185.
Meir, P., Kruijt, B., Broadmeadow, M., Barbosa, E., Kull, O., Carswell, F., Nobre, A. and Jarvis, P.G. (2002) Acclimation of photosynthetic capacity to irradiance in tree canopies in relation to leaf nitrogen concentration and leaf mass per unit area. *Plant, Cell & Environment,* 25 (3), pp. 343-357.
Moore, S., Adu-Bredu, S., Duah-Gyamfi, A., Addo-Danso, S.,D., Ibrahim, F., Mbou, A.T., De Grandcourt, A., Valentini, R., Nicolini, G. and Djagbletey, G. (2018) Forest biomass, productivity and carbon cycling along a rainfall gradient in West Africa. *Global Change Biology,* 24 (2), pp. e496-e510.
Moreno, J.F., Asner, G.P., Bach, H., Belenguer, T., Kneubühler, M. and Schaepman, M.E. (2006) Fluorescence explorer (FLEX): An optimised payload to map vegetation photosynthesis from space.
Natale, V.G., Kafri, A., Tidhar, G.A., Chen, M., Feingersh, T., Sagi, E., Cisbani, A., Baroni, M., Labate, D. and Nadler, R. (2013) SHALOM—Space-borne hyperspectral applicative land and ocean mission. IEEE, pp. 1.
Ordoñez, J.,C., Van Bodegom, P.,M., Witte, J.M., Wright, I.J., Reich, P.B. and Aerts, R. (2009) A global study of relationships between leaf traits, climate and soil measures of nutrient fertility. *Global Ecology and Biogeography,* 18 (2), pp. 137-149.
Pajares, G. (2015) Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs). *Photogrammetric Engineering & Remote Sensing,* 81 (4), pp. 281-330.
Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L. and Canadell, J.G. (2011) A large and persistent carbon sink in the world's forests. *Science,* 333 (6045), pp. 988-993.
Pignatti, S., Palombo, A., Pascucci, S., Romano, F., Santini, F., Simoniello, T., Umberto, A., Vincenzo, C., Acito, N. and Diani, M. (2013) The PRISMA hyperspectral mission: Science activities and opportunities for agriculture and land monitoring. IEEE, pp. 4558.
Poorter, L. and Bongers, F. (2006) Leaf traits are good predictors of plant performance across 53 rain forest species. *Ecology,* 87 (7), pp. 1733-1743.

Poorter, L. and Markesteijn, L. (2008) Seedling traits determine drought tolerance of tropical tree species. *Biotropica,* 40 (3), pp. 321-331.
Porder, S., Asner, G.P. and Vitousek, P.M. (2005) Ground-based and remotely sensed nutrient availability across a tropical landscape. *Proceedings of the National Academy of Sciences,* 102 (31), pp. 10909-10912.
Reich, P.B., Walters, M.B. and Ellsworth, D.S. (1997) From tropics to tundra: global convergence in plant functioning. *Proceedings of the National Academy of Sciences,* 94 (25), pp. 13730-13734.
Riipi, M., Haukioja, E., Lempa, K., Ossipov, V., Ossipova, S. and Pihlaja, K. (2004) Ranking of individual mountain birch trees in terms of leaf chemistry: seasonal and annual variation. *Chemoecology,* 14 (1), pp. 31-43.
Sakschewski, B., von Bloh, W., Boit, A., Rammig, A., Kattge, J., Poorter, L., Peñuelas, J. and Thonicke, K. (2015) Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. *Global Change Biology,* 21 (7), pp. 2711-2725.
Santiago, L.S., Goldstein, G., Meinzer, F.C., Fisher, J.B., Machado, K., Woodruff, D. and Jones, T. (2004a) Leaf photosynthetic traits scale with hydraulic conductivity and wood density in Panamanian forest canopy trees. *Oecologia,* 140 (4), pp. 543-550.
Santiago, L.S., Kitajima, K., Wright, S.J. and Mulkey, S.S. (2004b) Coordinated changes in photosynthesis, water relations and leaf nutritional traits of canopy trees along a precipitation gradient in lowland tropical forest. *Oecologia,* 139 (4), pp. 495-502.
Santiago, L.S. and Wright, S.J. (2007) Leaf functional traits of tropical forest plants in relation to growth form. *Functional Ecology,* 21 (1), pp. 19-27.
Scheiter, S., Langan, L. and Higgins, S.I. (2013) Next-generation dynamic global vegetation models: learning from community ecology. *New Phytologist,* 198 (3), pp. 957-969.
Skidmore, A.K., Ferwerda, J.G., Mutanga, O., Van Wieren, S.,E., Peel, M., Grant, R.C., Prins, H.H.T., Balcik, F.B. and Venus, V. (2010) Forage quality of savannas—Simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery. *Remote Sensing of Environment,* 114 (1), pp. 64-72.
Smith, G.M. and Milton, E.J. (1999) The use of the empirical line method to calibrate remotely sensed data to reflectance. *International Journal of Remote Sensing,* 20 (13), pp. 2653-2662.
Smith, M., Martin, M.E., Plourde, L. and 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 (6), pp. 1332-1337.
Stein, B.R., Thomas, V.A., Lorentz, L.J. and Strahm, B.D. (2014) Predicting macronutrient concentrations from loblolly pine leaf reflectance across local and regional scales. *GIScience & Remote Sensing,* 51 (3), pp. 269-287.
Sterck, F.J., Poorter, L. and Schieving, F. (2006) Leaf traits determine the growth-survival trade-off across rain forest tree species. *The American Naturalist,* 167 (5), pp. 758-765.
Stuffler, T., Kaufmann, C., Hofer, S., Förster, K., Schreier, G., Mueller, A., Eckardt, A., Bach, H., Penné, B. and Benz, U. (2007) The EnMAP hyperspectral imager—An advanced optical payload for future applications in Earth observation programmes. *Acta Astronautica,* 61 (1-6), pp. 115-120.
Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., Hünniger, D., Bartholomeus, H., Becker, R. and Kooistra, L. (2014) A lightweight hyperspectral

mapping system and photogrammetric processing chain for unmanned aerial vehicles. *Remote Sensing,* 6 (11), pp. 11013-11030.
Tiessen, H., Hauffe, H. and Mermut, A.R. (1991) Deposition of Harmattan dust and its influence on base saturation of soils in northern Ghana. *Geoderma,* 49 (3-4), pp. 285-299.
van de Weg, M.,J., Meir, P., Grace, J. and Atkin, O.K. (2009) Altitudinal variation in leaf mass per unit area, leaf tissue density and foliar nitrogen and phosphorus content along an Amazon-Andes gradient in Peru. *Plant Ecology & Diversity,* 2 (3), pp. 243-254.
Verrelst, J., Malenovský, Z., Van der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J., Lewis, P., North, P. and Moreno, J. (2019) Quantifying vegetation biophysical variables from imaging spectroscopy data: a review on retrieval methods. *Surveys in Geophysics,* 40 (3), pp. 589-629.
Webb, C.T., Hoeting, J.A., Ames, G.M., Pyne, M.I. and LeRoy Poff, N. (2010) A structured and dynamic framework to advance traits-based theory and prediction in ecology. *Ecology Letters,* 13 (3), pp. 267-283.
Weier, J. and Herring, D. (2000) Measuring Vegetation (NDVI & EVI). NASA Earth Observatory. *Washington, DC, USA,* .
West, J.S., Bravo, C., Oberti, R., Lemaire, D., Moshou, D. and McCartney, H.A. (2003) The potential of optical canopy measurement for targeted control of field crop diseases. *Annual Review of Phytopathology,* 41 (1), pp. 593-614.
Wold, S., Ruhe, A., Wold, H. and Dunn, I., W.J. (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing,* 5 (3), pp. 735-743.
Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C. and Diemer, M. (2004) The worldwide leaf economics spectrum. *Nature,* 428 (6985), pp. 821-827.
Wu, J., Chavana-Bryant, C., Prohaska, N., Serbin, S.P., Guan, K., Albert, L.P., Yang, X., van Leeuwen, W.,J.D., Garnello, A.J. and Martins, G. (2017) Convergence in relationships between leaf traits, spectra and age across diverse canopy environments and two contrasting tropical forests. *New Phytologist,* 214 (3), pp. 1033-1048.
Zhang, L., Sun, X., Wu, T. and Zhang, H. (2015) An analysis of shadow effects on spectral vegetation indexes using a ground-based imaging spectrometer. *IEEE Geoscience and Remote Sensing Letters,* 12 (11), pp. 2188-2192.

6.12. Supplementary material

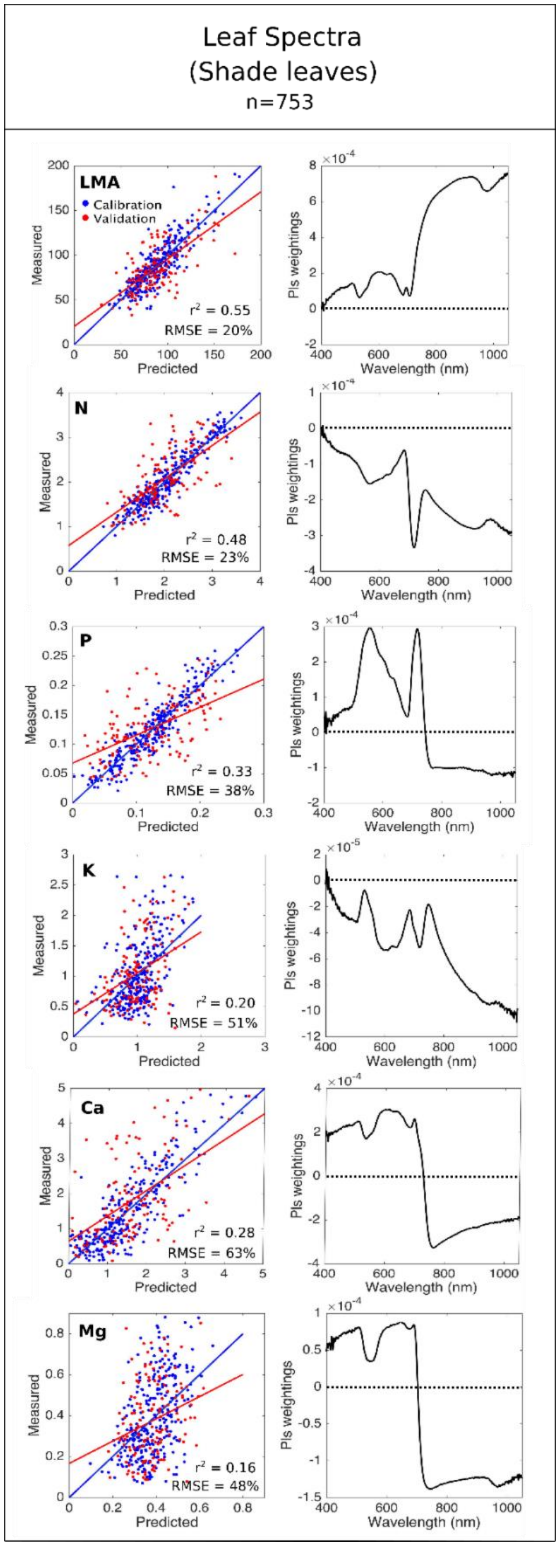


Figure S6.1: (left) Relationships between measured and predicted leaf traits using shade leaf reflectance spectra from Figure 2 and PLS regression analysis. n = 753. Blue line = r^2-cal regression line. Red line = r^2-val regression line. RMSE and r^2-cal values displayed on graph. (right) PLSR spectral weightings for each leaf trait. Deviation from 0 indicates the parts of the spectrum that most strongly influence the empirical model.

Chapter 7:

Mapping photosynthetic capacity in a Bornean tropical forest from airborne and satellite imagery: challenges and recommendations



7.1 Chapter Statement

This chapter uses aircraft-collected hyperspectral imagery and Sentinel-2 imagery to develop predictive models of photosynthetic capacity along a forest disturbance gradient in Malaysian Borneo. Building on the results in **Chapter 2 & 3**, it expands a focus from mapping biochemical leaf traits to mapping process-based traits: $Vcmax_{25}$ and $Jmax_{25}$.

With increased spatial coverage of spectral imagery, it explores some of key data quality issues affecting optical imagery, including varying illumination conditions during airborne hyperspectral collection, and the cloud-masking and surface reflectance accuracy of Sentinel-2's atmospheric correction processor. It highlights current limitations of optical remote sensing and presents best-practice recommendations for future airborne imaging studies.

7.2 Author Contributions

The ideas, research questions, study design, data processing, data analysis and writing presented in this chapter are the work of the first author, Eleanor Thomson. These activities greatly benefited from the contributions of Marcus Spiegel and Yadvinder Malhi.

Leaf photosynthesis collection was carried out by a team led by Sabine Both, with assistance from Terhi Riutta. Airborne hyperspectral imagery was collected by the Natural Environmental Research Council's Airborne Research Facility (NERC ARF), in collaboration with David Coomes and his research group at Cambridge University. Tree crown maps were created by Martin Svátek.

7.3. Abstract

Plant photosynthetic capacity is determined by the maximum carboxylation rate ($Vcmax_{25}$) and the maximum electron transport rate ($Jmax_{25}$) at 25°C. Landscape level maps of photosynthetic capacity would help capture potentially important spatial and seasonal variation for use in empirical and modelling studies. Accurate estimates of photosynthetic capacity are particularly important in tropical forests, which make an outsized contribution to terrestrial productivity. Imaging spectroscopy has emerged as a promising technique for estimating photosynthetic capacity at the leaf level. In this study, we present evidence that airborne imagery spectroscopy can be used to predict $Vcmax_{25}$ and $Jmax_{25}$ at the canopy scale. We combine aerial hyperspectral imagery (400-2500nm), and Sentinel-2 imagery (10 bands), with leaf photosynthesis measurements and partial least squares regression (PLSR) to characterize pixel-level variation in $Vcmax_{25}$ and $Jmax_{25}$ across a forest disturbance gradient in Malaysian Borneo. However, we find that imagery quality issues, many of which are exacerbated in the tropics, reduce modelling accuracy. We find large cross-sensor disagreements between aerial hyperspectral and Sentinel imagery collections, and large internal sensor disagreements between Sentinel images taken on different days. Exploring the causes of sensor disagreement, we find aerial hyperspectral imagery is highly sensitive to illumination conditions at time of capture, and Sentinel-2 imagery suffers from two main sources of processing error: (1) Sentinel's cloud detection algorithm (SCL), which is only 76% accurate over Bornean tropical forest (2) Sentinel-2 surface reflectance values, which are unstable over tropical forests, especially in the visible bands, likely due to issues arising from atmospheric aerosol and water vapour retrieval. We find that a Sentinel-2 time-series composite displays the most credible reflectance patterns and outperforms a single Sentinel scene in predicting $Vcmax_{25}$ and $Jmax_{25}$ (r^2=0.21 and 0.28). The strongest model performance however was produced by tuning the high spectral and spatial resolution of the aerial hyperspectral data with the improved signal-to-noise ratio of the Sentinel composite (r^2=0.36 and 0.43 for $Vcmax_{25}$ and $Jmax_{25}$ respectively), demonstrating the importance of diagnosing and minimising data quality issues. Throughout the paper, we highlight current limitations of optical remote sensing and present best-practice recommendations for future airborne imaging studies.

7.4. Introduction

Tropical forests are one of the most productive ecosystems on Earth. They cover 7% of the Earth's land surface and account for one-third of terrestrial productivity (Beer *et al.* 2010; Malhi *et al.* 2012). Yet sparse ground data results in large uncertainties to this estimate and may fail to capture important spatial and seasonal variation.

Terrestrial biosphere models (TBMs) are the principal means for generating regional and global projections of terrestrial carbon cycling. Two main leaf traits ($Vcmax_{25}$ and $Jmax_{25}$) are used to parametrize empirical models of photosynthesis, including the widely-adopted Farquhar-von Caemmerer—Berry model (Farquhar *et al*. 1980). $Vcmax_{25}$ represents light-saturated conditions at 25°C, where photosynthesis is limited by the maximum carboxylation rate of RuBisCo, while $Jmax_{25}$ represents the light-limited state at 25°C, where photosynthesis is limited by electron transport. In TBMs, these trait data are typically sourced from the literature, or public databases, and partitioned into five to sixteen discrete plant functional types (PFTs) to describe variation across the entire planet (Rogers, 2014; Wullschlager *et al*. 2014). Tropical forests are typically assigned a single value of each of these photosynthetic traits, or values are separated into tropical evergreen and tropical deciduous forests. Such generalisation fails to capture the magnitude, regional variation, and seasonality of ecosystem photosynthetic potential (Zaehle *et al*. 2005; Bonan *et al.* 2011), impeding our ability to predict the photosynthetic response of tropical forests to a changing environment.

Remote sensing would be an ideal method to generate spatially continuous photosynthetic data. At the leaf level, numerous studies have used hyperspectral reflectance data (400-2500 nm) and partial least squares regression (PLSR) to develop predictive models of biochemical leaf traits, including light-absorbing pigments, nutrient concentrations, and leaf mass per area (Asner and Martin *et al*. 2008; Serbin *et al*. 2014; Asner *et al*. 2015; Doughty *et al*. 2017; Thomson *et al*. 2018). A few studies have attempted to measure photosynthetic capacity directly using this method, and their results have been promising. These studies are based on the premise that photosynthetic processes are influenced by physical leaf characteristics, such as chlorophyll and nitrogen concentration, which in turn affect the optical properties of vegetation. Highest

prediction accuracies of $Vcmax_{25}$ and $Jmax_{25}$ (r^2 = 0.88-0.92) have been achieved in crop studies (Serbin *et al.* 2012; Ainsworth *et al*. 2014; Heckmann *et al*. 2017). Across forest ecosystems, Doughty *et al*. (2011) obtained lower prediction accuracies of $Vcmax_{25}$ and $Jmax_{25}$ ($r^2 \leq 0.52$) for 11 tropical tree species, while Dechant *et al*. (2017), achieved higher estimates (r^2 = 0.64-70) for 37 temperate deciduous tree species. The best photosynthetic capacity results (r^2 = 0.90) were reported by Wu *et al*. (2019) for tropical forest leaves measured across Panama and Brazil.

Scaling from the plot to the canopy level, Serbin *et al*. (2015) used high-altitude Airborne Visible/Infrared Imaging Spectroscopy (AVIRIS) imagery to achieve very good prediction results (r^2 = 0.90) of $Vcmax_{25}$ for nine crop species. Similar studies have used hyperspectral aircraft imagery to map biochemical leaf traits across large swathes of tropical forests (e.g. Asner *et al*. 2016; Asner *et al*. 2017; Swinfield *et al*. 2018). Yet to the best of our knowledge, no study to date has attempted to use hyperspectral aircraft imagery to retrieve photosynthetic information over a tropical forest.

While maintaining high spatial and spectral resolution, aircraft imaging studies generally represent a single snapshot in time, precluding their use for temporal analyses. In contrast, most satellite imagery is captured at a lower spatial and spectral resolution, but at frequent intervals. The Sentinel-2 mission of the European Space Agency's (ESA) Copernicus programme currently represents the highest resolution, openly-accessible satellite imagery, with a ground sampling distance of 10m in the VNIR (20 m in the SWIR), 13 spectral bands and a revisit time of 5 days. Despite capturing fewer than 10% of the spectral bands measured by a hyperspectral sensor, the majority of Sentinel-2 spectral channels are clustered in the visible, red-edge and near-infrared parts of the spectrum - regions that contain many absorption features linked to biochemical leaf traits and are thus disproportionally important in measuring these traits (Curran, 1989; Kokaly *et al*. 2009; Ustin *et al*. 2009). Accordingly, Sentinel-2 imagery has been used to map a wide range of vegetation traits across the tropics (Aguirre-Gutiérrez *et al*. 2021), in regional temperate forests (Gara *et al*. 2019; Ma *et al*. 2019) and in the Arctic tundra (Thomson *et al*. 2021). As leaf-level models of photosynthesis display the largest model coefficients in the same spectral regions used to predict nitrogen concentration, leaf water content and internal leaf structure (Serbin *et al*. 2012; Ainsworth *et al*. 2014; Meacham-Hensold *et al*.

2019), these results suggest good potential for using Sentinel-2 imagery to measure and map photosynthetic capacity.

A key limitation to imaging studies however, is the vulnerability of optical imagery to data quality issues. Aircraft and satellite imagery are particularly sensitive to illumination and atmospheric conditions, such as sun angle, bidirectional reflectance (BRDF) effects, cloud cover, water vapour concentration, and the presence of smoke and haze. As predictive trait models rely on small spectral differences to determine trait values, corrupt spectral data may be propagated through the model to produce inaccuracies in the pattern or magnitude of trait values. These inaccuracies may be particularly pronounced when reflectance errors are not spatially or temporally consistent.

To correct for some of these effects in Sentinel-2 imagery, the atmospheric correction processor Sen2Cor was developed by Telespazio VEGA Deutschland GmbH on behalf of ESA (Main-Knorn *et al*. 2017). The main purpose of Sen2Cor is to perform atmospheric, terrain and cirrus correction on Sentinel-2 Level-1C Top-Of-Atmosphere images and deliver a reliable Level-2A Bottom-Of-Atmosphere reflectance product. In the process, it produces an Aerosol Optical Thickness (AOT) product, a Water Vapour (WV) product, and a Scene Classification Layer (SCL) that detects pixels covered by cloud or cloud shadow. Sen2Cor is ESA's official surface reflectance processing product and Level-2A imagery is integrated into the Copernicus Open Access Hub, the Sentinel Playground and Google Earth Engine. Results have been independently validated and shown to have good agreement with surface reflectance values (r^2=>0.9: Li *et al*. 2018). A comparison of the SCL with similar cloud-detection products, has shown it outperform its competitors (e.g. Zekoll *et al*. 2021), or comes a close second to Fmask4 (Sanchez *et al*. 2020).

In contrast to satellite instruments, aircraft-mounted sensors are typically flown below the level of the clouds, reducing issues associated with cloud cover. However, variable illumination conditions arising from cloud shadow and ground haze can still affect imagery quality. These issues are particularly challenging in humid tropical forests, which are some of the cloudiest regions in the world and where the level of haze from burning biomass has increased rapidly due to human and climate drivers (Streets *et al*. 2009; Eck *et al*. 2013; Chuvieco *et al*. 2021). To reduce cloud and aerosol effects, aircraft

are often flown in the morning when the atmosphere is clearest, but this trade-off reduces atmospheric interference, while increasing illumination issues associated with low sun angle and BRDF effects, especially across a structurally highly complex forest canopy.

This paper sets out to develop the strongest possible model of photosynthetic capacity across a forest disturbance gradient in Malaysian Borneo using two sources of airborne imagery: aerial hyperspectral (400-2500 nm) and Sentinel-2 imagery (10 bands). Given the vulnerability of remote sensing data to quality issues, we first evaluate the shortcomings of the imagery by (1) checking for consistency across the two sensors, and (2) checking for consistency within the Sentinel product, when multiple collection dates are available. We investigate the most likely causes of the inconsistencies observed. Based on an understanding of the strengths and limitations of the two datasets, we tune the aerial hyperspectral imagery using a Sentinel time-series composite. Combining the high spectral and spatial resolution of the airborne hyperspectral data with the stronger signal-to-noise ratio of the Sentinel time-series composite produced the most accurate model of both $Vcmax_{25}$ and $Jmax_{25}$.

Specifically, this paper asks:

1. What are the most likely causes of cross-sensor disagreements between airborne hyperspectral imagery and Sentinel imagery taken in the same season?
2. What are the most likely causes of internal sensor disagreements between three single Sentinel images, and a Sentinel time-series composite, taken in the same season?

    i)Inaccuracy in Sentinel's cloud detection algorithm (SCL)?

    ii)Inaccurate surface reflectance values?

3. Which type of spectral imagery produces the strongest PLSR model for estimating photosynthetic capacity ($Vcmax_{25}$ and $Jmax_{25}$) across a forest disturbance gradience in Malaysian Borneo?
4. Does combining the high spectral and spatial resolution of the airborne hyperspectral imagery with the stronger signal-to-noise ratio of a Sentinel time-series composite improve estimation of $Vcmax_{25}$ and $Jmax_{25}$?

7.5. Materials and Methods

Study Sites

Data were collected across seven, one-hectare tropical forests plots in the Malaysian state of Sabah, northern Borneo *(***Figure 7.1, Table 7.1**). The plots span a forest disturbance gradient from old-growth primary forest to logged forest fragments (Riutta *et al*. 2018) and form part of the Global Ecosystems Monitoring (GEM) network (Malhi *et al*. 2021). Four of the plots are located in old-growth forest in the Maliau Basin and the Danum Valley Conservation Area. Although some low intensity logging occurred around these sites in the 1970s, both regions remain mainly undisturbed and are dominated by tall old-growth dipterocarp forest. The three remaining plots are located in the Kalabakan Reserve and form part of the Stability of Altered Forest Ecosystems project (SAFE, www.safeproject.net; Ewers *et al.* 2011). These plots were first logged in the mid-1970s, with repeated rounds of logging taking place during 1990-2008. Two plots (B North and B South) were selectively logged four times, while the third plot (E) was selectively logged twice. Most of the forest surrounding these forest fragments has been entirely converted to oil palm.

All three areas are part of the Yayasan Sabah Forest Management Area and belong to a formerly connected area of evergreen dipterocarp rainforest characterised by high species richness and many tall, emergent trees. The region has a moist tropical climate with an annual daily mean temperature of 26.7 °C and annual precipitation of approximately 2600-2700 mm (Walsh and Newbery, 1999). While precipitation is high all year round, significantly higher rainfall occurs during the northerly monsoon period of October to February, with a drier season spanning March to September. Occasional droughts are associated with El Niño Southern Oscillation events, with the last El Niño drought affecting Sabah from 2015-16 (Rifai *et al*. 2019). The forest soils across Sabah are mostly orthic Acrisols or Ultisols over undulating clay (Marsh and Greer, 1992; **Table 7.1**).

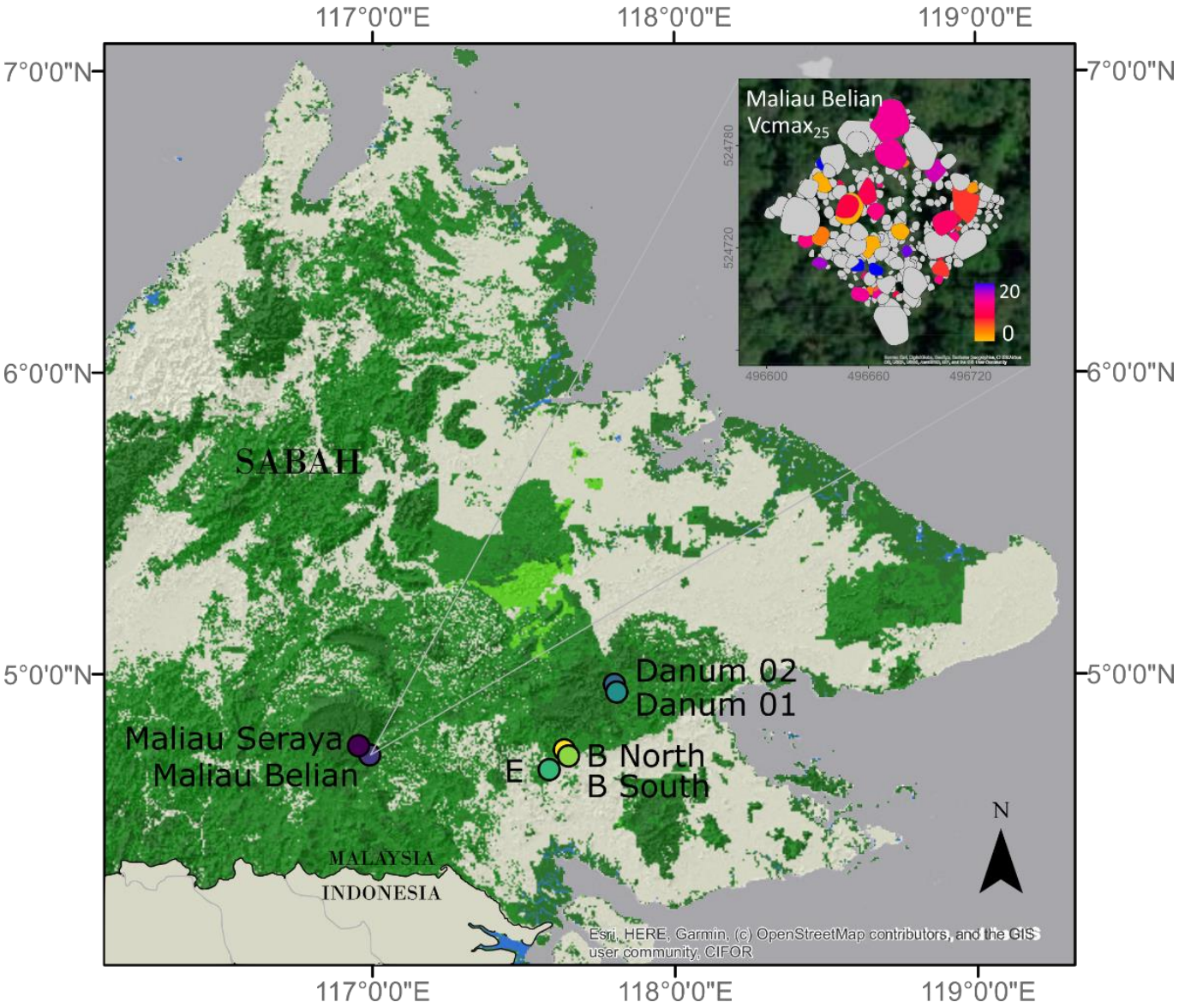


Figure 7.1: A map of forest cover across the Malaysian state of Sabah, northern Borneo, and location of the seven GEM plots. Inset: Geolocated tree crown map for Maliau Belian (MLA-01), created used a ground-based laser method (. Sampled tree crowns are coloured with their corresponding field-measured $Vcmax_{25}$ measurements ($\mu mol/m^{-2}$ s^{-1}).

Plot Name and Code	Plot Location	Disturbance	Soil type and topography	Basal area of trees >10 cm DBH (m^2)	Mean canopy gap fraction (%)	Big trees (DBH >50 cm)	Pioneer tree species (% of basal area)
B South (SAF-01)	SAFE Project (4.732°, 117.619°)	Heavily logged	Clay; mostly flat with a moderate slope on one edge	6.81 ± 1.00	15.0 ± 2.14	1	28.1 ± 4.3
B North (SAF-02)	SAFE Project (4.739°, 117.617°)	Heavily logged	Clay; undulating	11.1 ± 1.81	13.4 ± 1.51	0	57.2 ± 5.8
E (SAF-03)	SAFE Project (4.691°, 117.588°)	Moderately logged	Clay; steep slope	19.6 ± 1.88	11.2 ± 1.29	10	21.5 ± 5.0
Danum 01 (DAN-04)	Danum Valley Conservation Area (4.951°, 117.796°)	Old-growth	Clay; steep slope	32.0 ± 3.30	10.5 ± 1.00	34	0.7 ± 1.0
Danum 02 (DAN-05)	Danum Valley Conservation Area (4.953°, 117.793°)	Old-growth	Clay; flat	30.6 ± 3.37	11.3 ± 1.5	26	0.1 ± 0.0
Maliau Belian (MLA-01)	Maliau Basin Conservation Area (4.747°, 116.970°)	Old-growth	Clay; undulating	41.6 ± 3.59	8.70 ± 0.62	47	0.2 ± 0.3
Maliau Seraya (MLA-02)	Maliau Basin Conservation Area (4.754°, 116.950°)	Old-growth	Clay; moderate slope	34.7 ± 2.74	7.04 ± 0.55	56	1.7 ± 1.2

Table 7.1: Characteristics for seven GEM plots in Sabah, Malaysian Borneo. The logged plots are ranked from most to least intensively logged. Plot mean ±1 SE. Adapted from Riutta *et al.* (2018).

Leaf photosynthesis collection

Leaf photosynthesis measurements were carried out by Both *et al*. (2019) between July-December, 2015. A_s, A_{max} and dark respiration (R_{dark}) measurements were taken across all seven plots using a Licor 6400XT. A combination of the largest and most numerous species were measured, that together represented >70% of the basal area of each plot. In 57% of cases, only one individual per species was present in the plot, otherwise, the 1-3 individuals with the highest DBH were sampled. Canopy sun leaves were collected by tree climbing. Three fully expanded, undamaged leaves were randomly sampled per branch, with 30 photosynthesis measurements averaged per leaf. Measurements were checked for consistency and those displaying anomalous light or CO_2 recordings removed. See Both *et al*. (2019) for more details.

A-Ci-curves frequently used to measure $Vcmax_{25}$ and $Jmax_{25}$ were not taken in any of the plots, as analysis of photosynthetic capacity is time-consuming and fell outside the scope of Both *et al*. (2019). In order to derive $Vcmax_{25}$ and $Jmax_{25}$ information, the one-point

method from the plantecophys R package (Duursma, 2015) was used. This method generates $Vcmax_{25}$ and $Jmax_{25}$ estimates by inverting the Farquhar photosynthesis equation, and is based on a study by de Kauwe *et al*. (2016) who used A_{sat} and R_{day} measurements to estimate $Vcmax_{25}$ with an r^2 of 0.95 and RMSE of 17.1 μmol m^{-2} s^{-1}. We independently tested the one-point method at GEM sites where *A*-Ci-curves were collected: Australia, Brazil and Gabon. Good agreement was found between values derived from A-Ci-curves and those estimated using the one-point method, with r^2 falling between 0.57-79 for $Vcmax_{25}$ and 0.63-92 for $Jmax_{25}$ (**Figure S7.1**).

Airborne hyperspectral campaign

Airborne hyperspectral imagery was collected across Sabah by the Natural Environmental Research Council (NERC) Airborne Research Facility (ARF) in November 2014. An AisaFENIX sensor was used to capture visible to short-wave infrared data (380–2,500 nm) using two parallel spectrometers, mounted on a Dornier 228-201. The aircraft was flown at an altitude of 1,400–2,400 m a.s.l. with a flight speed of 120–140 knots. Spatial resolution varied between 1.5 - 3.5 m over the GEM plots, depending on survey altitude. At the lowest altitudes, the spectral resolution was reduced from an average of 6 nm to 12 nm, in order to maximize signal-to-noise ratio at the sensor. During imagery capture, all flights were connected to a ground-based Leica base station to allow for sub-meter accuracy georeferencing of the imagery.

Raw spectra were radiometrically calibrated to irradiance by NERC ARF using laboratory-collected informed radiometric calibration coefficients and spectral calibration. Signal-to-noise ratio was enhanced by averaging every three bands in the visible and the near-infrared regions (420–1,100 nm) and every two bands in the shortwave infrared region (1,100–2,400 nm). Images from each flight were georeferenced to a LiDAR-derived digital elevation model generated from the same flight. Atmospheric and Bidirectional Reflectance Distribution Function correction was applied with the rugged terrain algorithm in the Modtran radiative transfer model in ATCOR v. 6.3.2, to account for the solar and viewing angles of each pixel during flight and calculate ground reflectance. See Swinfield *et al*. (2018) for more details.

Hyperspectral imagery processing

Post NERC ARF processing, the imagery was filtered to remove pixels with an NDVI of <0.8 and a solar zenith angle of >60°. As many of the flights were conducted in the morning, when skies were clearest, the angle between the sun and the sensor was often large, resulting in a large proportion of poor-quality spectra. Remaining pixels with a canopy height of less than 4 m, as measured by the LiDAR canopy map, were also filtered to exclude non-tree vegetation.

For spectral analyses, the low spectral resolution imagery (190 bands) was resampled to match the high spectral resolution imagery (246 bands) by linearly interpolating between wavelengths. The spectra were trimmed at the far edges of each spectrometer and within the water absorption bands (1,350–1,480 nm and 1780–2032 nm), resulting in a final range of 400-900 nm and 1000-2400 nm.

Average hyperspectral curves over each GEM plot were computed using the remaining spectra pixels and GPS-located plot corner coordinates.

Sentinel-2 data

Sentinel-2 Bottom-Of-Atmosphere reflectance imagery from the 9th September, 18th October and 27th December, 2019 was downloaded from the Copernicus Open Access hub. These dates were chosen as they represented the first year that Level-2A products were available over Borneo, and were the closest seasonal dates to the November hyperspectral imagery collection. Within that, they were the only dates when 6-7 of the plots exhibited clear vegetation with no cloud cover (according to the SCL map).

A Sentinel-2A time-series composite was generated using Google Earth Engine. The composite was created by taking all the available Level-2 images over the plots between 1st September and 30th December 2019, filtering each image to leave only clear vegetation pixels (as classified by the image's corresponding SCL map), and taking the median value of all the remaining reflectance values for each pixel in the time-series.

To provide an accurate comparison to the hyperspectral imagery, the Sentinel images were resampled to the same spatial resolution as the hyperspectral imagery using bilinear interpolation. They were then masked by the hyperspectral imagery, in order to remove exactly the same pixels that had been filtered in the hyperspectral cleaning process.

Average Sentinel reflectance curves over each GEM plot were computed using the remaining spectra pixels and the same GPS-located plot corner coordinates. Bands 1, 9 and 10 were omitted from the Sentinel analyses, as these bands are not designed for ecological mapping.

While top-of-atmosphere Level-1C Sentinel imagery was available closer in time to the hyperspectral and photosynthesis data collection (earliest year: 2017), these images require atmospheric correction using the Sen2Cor processor, which necessitates considerable computing storage and resources. Thus, ready-computed Level-2A products were used instead. It is expected that the composition and structure of an undisturbed evergreen tropical forest will not change substantially in 4-5 years. During this time, plot surveys show little change in species distribution or basal area (GEM, unpublished data).

Evaluation of airborne hyperspectral data

One hypothesis for the observed inconsistencies between the airborne hyperspectral and Sentinel imagery is that the radiometric calibration performed by NERC ARF did not adequately correct for differences in illumination conditions as the aircraft flew under and out of clouds. In order to test this hypothesis, we calculated how often the plots are covered in cloud during the start of an average wet season. From 2015-2019, all the Sentinel images taken in October or November were assessed in Google Earth for atmospheric conditions over each plot (n=245). The images were manually reviewed and assigned to one of four classes: clear vegetation, thin cloud, thick cloud or cloud shadow. Thin cloud was characterised by the visible presence of vegetation beneath the cloud cover and cloud shadow was easily discernible as dark patches cast by neighbouring clouds.

Evaluation of Sentinel-2 imagery

Evaluating the SCL

To evaluate the accuracy of Sen2Cor's SCL layer over the forests of northern Borneo, the median SCL classification class for each GEM plot was compared against a manual assessment in Google Earth Engine. All Sentinel images during 2019 were included in the analysis. The same classes were used as above: clear vegetation (SCL class 4), thin cloud (SCL class 10), thick cloud (SCL classes 8 and 9) and cloud shadow (SCL class 3).

Evaluating Sen2Cor's surface reflectance values

The accuracy of the Sen2Cor processer for calculating surface reflectance values over tropical forest were investigated by plotting a time-series of three Sentinel bands (Green (Band 3), Red (Band 4) and NIR (Band 8)) over the seven GEM plots. Only Sentinel images that had been manually categorised as exhibiting clear vegetation over the plot in question were included in the analysis. Reflectance values were only plotted over the dry season, as the rest of the year included too few clear sky images for a robust analysis. A rolling 5-day time series was plotted, as well as an average monthly mean.

Using the same method, two other GEM sites - located in Ghana and Brazil - were included in the analysis to compare the reliability of Sen2Cor's surface reflectance values over different types of tropical forest. The plots in Ghana (Thomson *et al*. 2018) span a precipitation gradient from wet evergreen tropical forest (Ankasa), to drier semi-deciduous tropical forest (Bobiri) to dry forest-savannah transition (Kogyae). The GEM plots in Brazil (Gvozdevaite *et al*. 2018) are located at the Amazon-Cerrado transition and represent tall closed canopy forest (VCR-02), transitional forest (NXV-02) and savanna forest (NXV-01).

In order to distinguish remote sensing artefacts from a possible authentic ecological signal, each plot and spectral band time series was fitted with a third degree polynomial curve of best fit. This analysis was based on the premise that phenological change typically exhibits a smooth time-series curve, often be mirrored by reflectance patterns

(e.g. MacBean *et al*. 2015; Cheng *et al*. 2020; Yang *et al*. 2020), whereas remote sensing artefacts weaken authentic trends by introducing random error. As our images covered approximately one season, we fitted the data with a smooth third-order polynomial curve that allowed for up to two changes in trend direction. How accurately the polynomial curve fitted the reflectance data was calculated using an adjusted R-squared, and the standard deviation of the residuals (RMSE). An adjusted R-squared was used to account for the varying number of clear vegetation images over each plot. The analysis was also repeated using a second-order polynomial curve with very similar results (**Table S7.2**).

Secondly, in order to distinguish remote sensing artefacts from a potential authentic ecological signal, we assumed that plots located close to each other would experience synchronous ecological changes, resulting in similar reflectance trends (Tobler's First Law of Geography, 1970), whereas the presence of random error would weaken this relationship. To test our data against this hypothesis, we paired plots based on their proximity to each other (<2.5km) and similar forest type. This grouping resulted in 7 pairs of plots: B North+South; Danum 01+02; Maliau Belian+Seraya; Ankasa 01+03; Bobiri SNR+Y53; Kogyae 02+03 and Nova Xavantina 01+02. For three spectral bands (Red, Green and NIR), the variance between the paired plots was computed as a proportion of the overall variance for that GEM site. These bands were chosen as they have been shown to disproportionately contain ecological information (Asner and Martin *et al*. 2008; Asner *et al*. 2015; Doughty *et al*. 2017) and are the bands most frequently used in vegetation indices. The same analysis was computed using the monthly mean data to smooth out any small reflectance variations within the plot pairs, with very similar results (**Table S7.3**).

Predicting foliar traits from hyperspectral imagery

Geolocated tree crown maps were used to match the spectral imagery with individual trees. Tree crown maps for each plot were created between September-October 2016, using the ground-based laser method described in Hédl *et al*. (2009) to map all stem and crown outlines to decimetre spatial precision (see Swinfield *et al*. (2018) for more information). Corresponding height measurements were used to filter overlapping tree crowns, leaving only the tallest sun-crowns that contributed the most to plot reflectance (**Figure 7.2A**).

Combining the tree crown maps and processed hyperspectral imagery, an average reflectance spectrum was computed for each tree crown. All pixels that fell wholly or partially within a tree crown were included in the average, but the spectral contribution of partial pixels was weighted by the percentage of the pixel that fell within the tree crown. To ensure a balance between high data quality and reasonable sample size, only tree crowns that were >50% covered in spectra pixels were included in the analysis. Analyses were repeated with tree crowns >30% and >80% covered in spectra pixels, but both approaches yielded lower results, due to reduced data quality or overly strict filtering respectively (**Table S7.5**).

Each tree spectrum was matched with the photosynthetic capacity data measured on that tree in the field (**Figure 7.2C**). For some trees, corresponding photosynthesis data was not available. For these individuals, representative trait values were attributed using the following method 1) individuals from which photosynthesis measurements had been taken kept their original trait values. 2) For individuals of a species from which a sampled individual was present in the plot, the trait value of the sampled individual was attributed. If more than one sampled individual of the same species was present in the plot, a distance-weighted mean was attributed. A distance-weighted mean was used to account for the influence of microclimates, micro-variations in soil, and the likelihood that individuals in close proximity share genetic information. 3) If individuals of a species were not sampled in the plot, but were sampled in a nearby plot, a species-mean from outside the plot was attributed. 4) For the old-growth Maliau and Danum plots, if individuals of a species were not sampled at that site, but were sampled at the other site,

a species-mean from that site was attributed. No trait values were assigned to species for which a sampled trait value did not exist. No trait values were trasmitted between old-growth and logged forests. Attributing species-levels means to unsampled individuals is supported by studies that show that taxonomy is more important than site in influencing trait distributions (Asner *et al.* 2014; Oliveras *et al.* 2020). In total, this method resulted in 222 matched spectra-trait tree crowns.

Partial least squares analysis was used to build predictive models of $Vcmax_{25}$ and $Jmax_{25}$ using the matched hyperspectral and photosynthesis data (Haaland and Thomas, 1989). The PLSR method is effective, as it uses the continuous spectrum as a single measurement, rather than carrying out a band-by-band analysis and reduces a large predictor matrix (246 spectral bands) down to a few relatively uncorrelated factors (latent variables). PLSR analyses were carried out using the PLSregress function in MATLAB. The PLSR models were validated using an unseen 30% of the dataset. Due to the random nature of the 70:30 split, 100 PLSR iterations were made for each spectra-trait analysis and evaluated using r^2 for the independent testing dataset and RMSE as a percentage of the sample mean.

Predicting foliar traits from Sentinel-2 imagery

To match photosynthesis traits values with Sentinel-2 imagery, the tree crown map with associated trait values was resampled to match the resolution and spatial extent of the Sentinel-2 pixels (**Figure 7.2D**). Only Sentinel pixels that were at least 60% covered by trait information were included in the analysis. The matched 10-band spectra-trait data were run through the same PLSR analyses described above. Models were built using two different Sentinel images – a single Sentinel image taken on 23rd September 2018, and the Sentinel composite product.

Tuning the hyperspectral imagery with a Sentinel composite

In order to combine the high spectral and spatial resolution of the airborne hyperspectral imagery with the stronger signal-to-noise ratio of the Sentinel time-series composite, the airborne hyperspectral imagery was 'tuned' using radiometric weights computed from

the offset between the airborne hyperspectral imagery and the Sentinel composite. Radiometric calibration weights were created for each plot using the average hyperspectral curve over that plot and the corresponding average Sentinel composite spectrum. The radiometric weights were applied to all hyperspectral pixels within that plot. The weights were created as follows: (1) At each of the 10 Sentinel wavelengths, the hyperspectral reflectance value at the nearest matching wavelength was subtracted from the Sentinel reflectance value to create a list of 10 offset factors. Eight out of ten of the Sentinel wavelengths fell within the first continuous section of hyperspectral spectra (from 400-900nm). For this section, the weights were linearly interpolated across the spectrum to create a list of 89 weights matching the 89 hyperspectral bands. (2) For the last three sections of spectra, less Sentinel information was available, so the weights were not interpolated. Instead, weights were created by taking the single nearest offset factor and applying it across the entire section of spectra. (3) Altogether, this method resulted in a list of 246 weights matching the 246 hyperspectral bands. For each plot, the plot-specific weights were added to all hyperspectral pixels within that plot, as well the average hyperspectral curve to view the result (**Figure 7.7**).

The tuned hyperspectral imagery was matched with the field-measured photosynthesis data and run through the same PLSR analyses described above.

All analyses were also carried out with A_{sat} and A_{max} values (see *7.10. Supplementary material* for results).

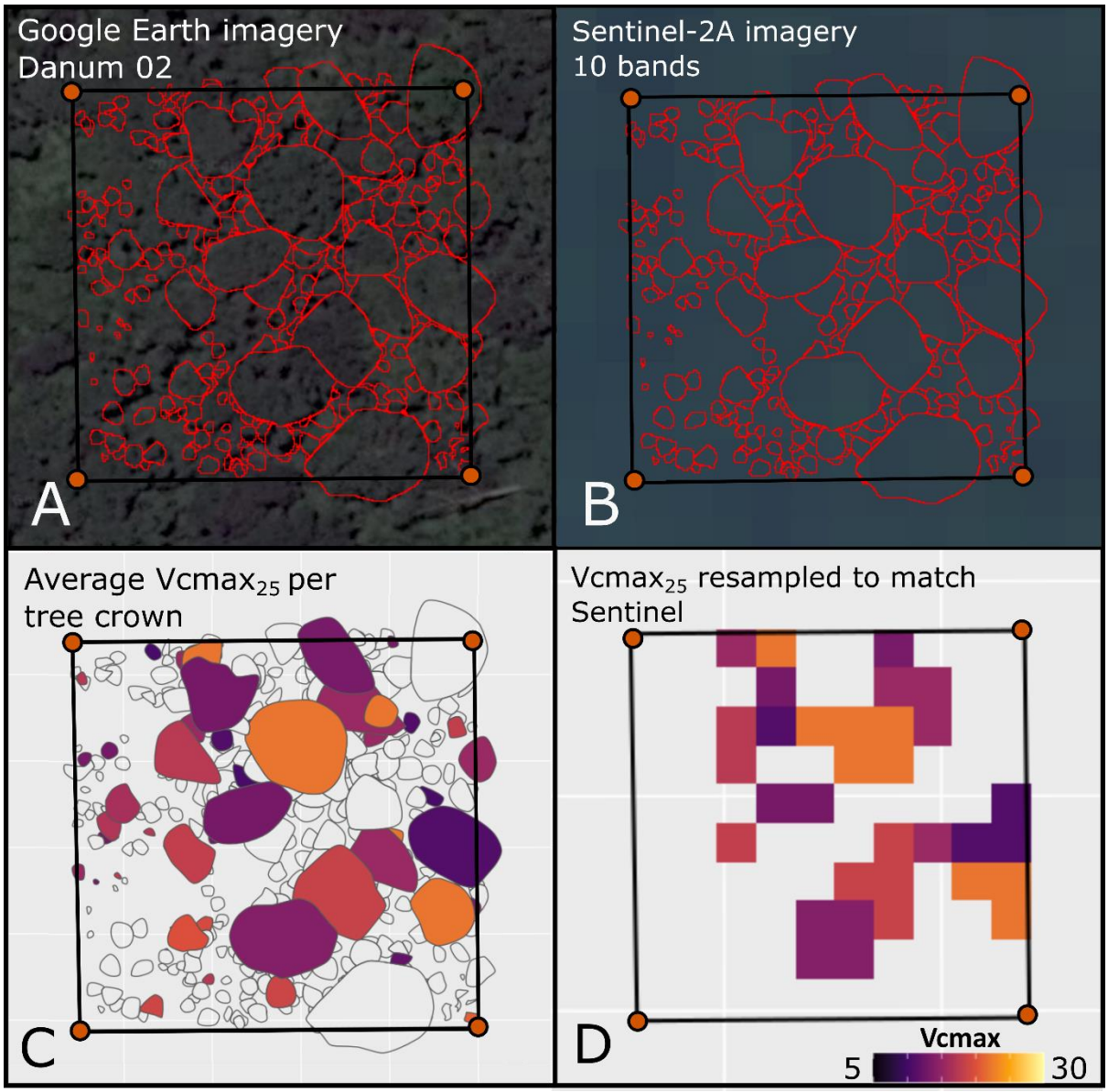


Figure 7.2: An example of the approach taken to match field measurements to Sentinel-2 imagery (A) Very high-resolution imagery of Danum 02 (taken from Google Earth: accessed 24/08/2021) overlain with plot corners and the laser tree crown map (B) Sentinel-2 imagery overlain with Danum 02 plot corners. (c) Tree crowns coloured with corresponding $Vcmax_{25}$ values. (D) $Vcmax_{25}$ values resampled to match resolution and extent of Sentinel-2 imagery. Only pixels >60% covered by tree crowns with trait information were included.

7.6 Results

Photosynthetic capacity results

Photosynthetic capacity values exhibited a clear forest disturbance gradient (**Figure 7.3**). $Vcmax_{25}$ and $Jmax_{25}$ means were significantly higher in the heavily logged plots of BNorth and BSouth, and lower in the old-growth forests of Danum and Maliau. The lightly logged plot (E), displayed means within the range of the old-growth plots, but with a long upper tail of values, spanning some of the highest values measured in BNorth and BSouth. Within the old-growth plots, Maliau Seraya displayed the highest $Vcmax_{25}$ values and Maliau Belian the lowest.

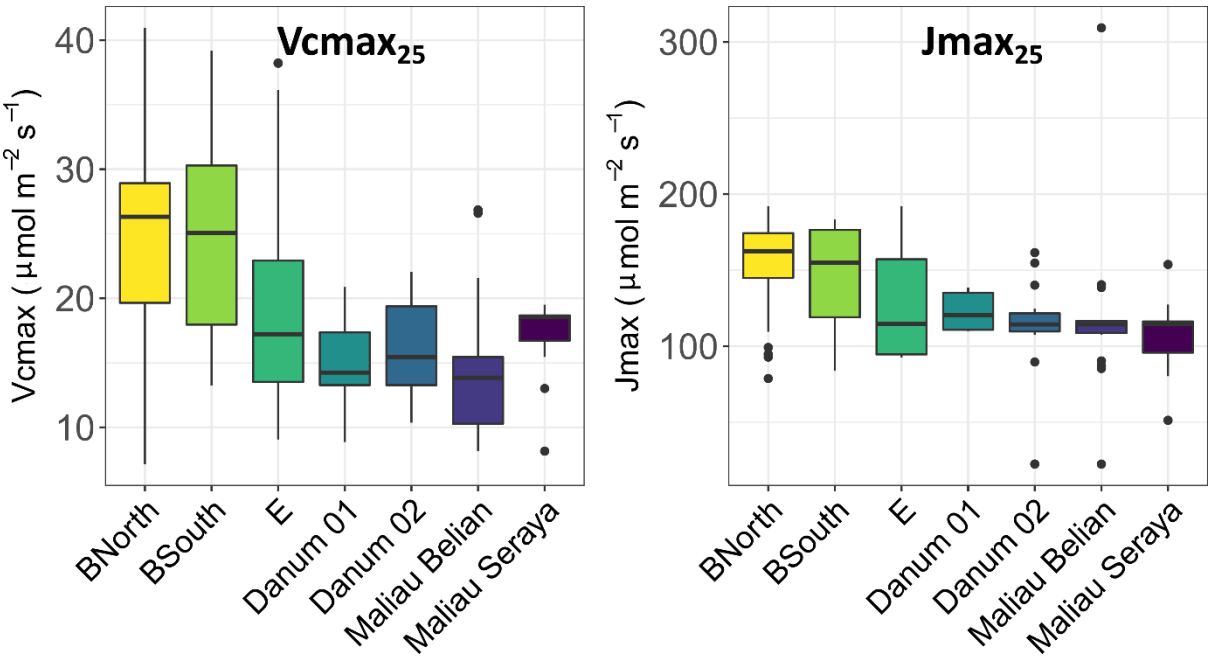


Figure 7.3: $Vcmax_{25}$ and $Jmax_{25}$ values across seven GEM plots in Borneo representing a forest disturbance gradient from logged to old-growth forest. Plots are ordered from heavily logged (BNorth and BSouth), to lightly logged (E), to old-growth forest (Danum and Maliau). $Vcmax_{25}$ and $Jmax_{25}$ values were derived using the one-point method (de Kauwe *et al.* 2016).

Spectral disagreements

Over the same plots, in the same season, there were large cross-sensor differences in reflectance patterns and magnitude between the airborne hyperspectral imagery collection and the Sentinel imagery collection. In addition, there were large internal sensor disagreements between Sentinel images taken on different days, and a Sentinel time-series composite of the same season (**Figure 7.4**).

Comparing the airborne imagery to the field-sampled $Vcmax_{25}$ and $Jmax_{25}$ results, hyperspectral reflectance patterns over the plots did not follow a clear forest disturbance gradient, although plots located close to each other exhibited similar reflectance values (**Figure 7.4(a)**). Danum 01, Danum 02 and E displayed extremely high reflectance in all wavelengths, peaking at a reflectance factor of 0.4. The Maliau plots displayed the lowest reflectance values across most wavelengths, peaking at a reflectance of just under 0.25. Hyperspectral curves for BNorth and BSouth were similar in magnitude to Maliau, except in the near-infrared where BNorth peaked at a slightly higher reflectance factor of 0.3.

The single Sentinel scenes displayed widely diverging spectral patterns (**Figure 7.4(c-e)**). Imagery captured on the 23rd of September, 2019 exhibited the most similar patterns to the Sentinel composite, although reflectance values for Danum 02 were significantly higher and reflectance values for E were significantly lower. Imagery captured on the 18th October, 2019 showed comparably high and low reflectance for Danum 02 and E, but otherwise no similar spectral patterns to September 23rd. Imagery captured on the 27th December, 2019 showed a tight clustering of spectral curves, with no comparison to the September or October imagery.

The Sentinel composite exhibited the most credible spectral patterns, demonstrating clear spectral separability between the logged and old-growth plots (**Figure 7.4(b)**). Highest reflectance was displayed by the three logged plots and lowest reflectance shown by the old-growth plots. Mirroring the $Vcmax_{25}$ results, Maliau Seraya displayed the higher reflectance of the old-growth plots and Maliau Belian the lowest.

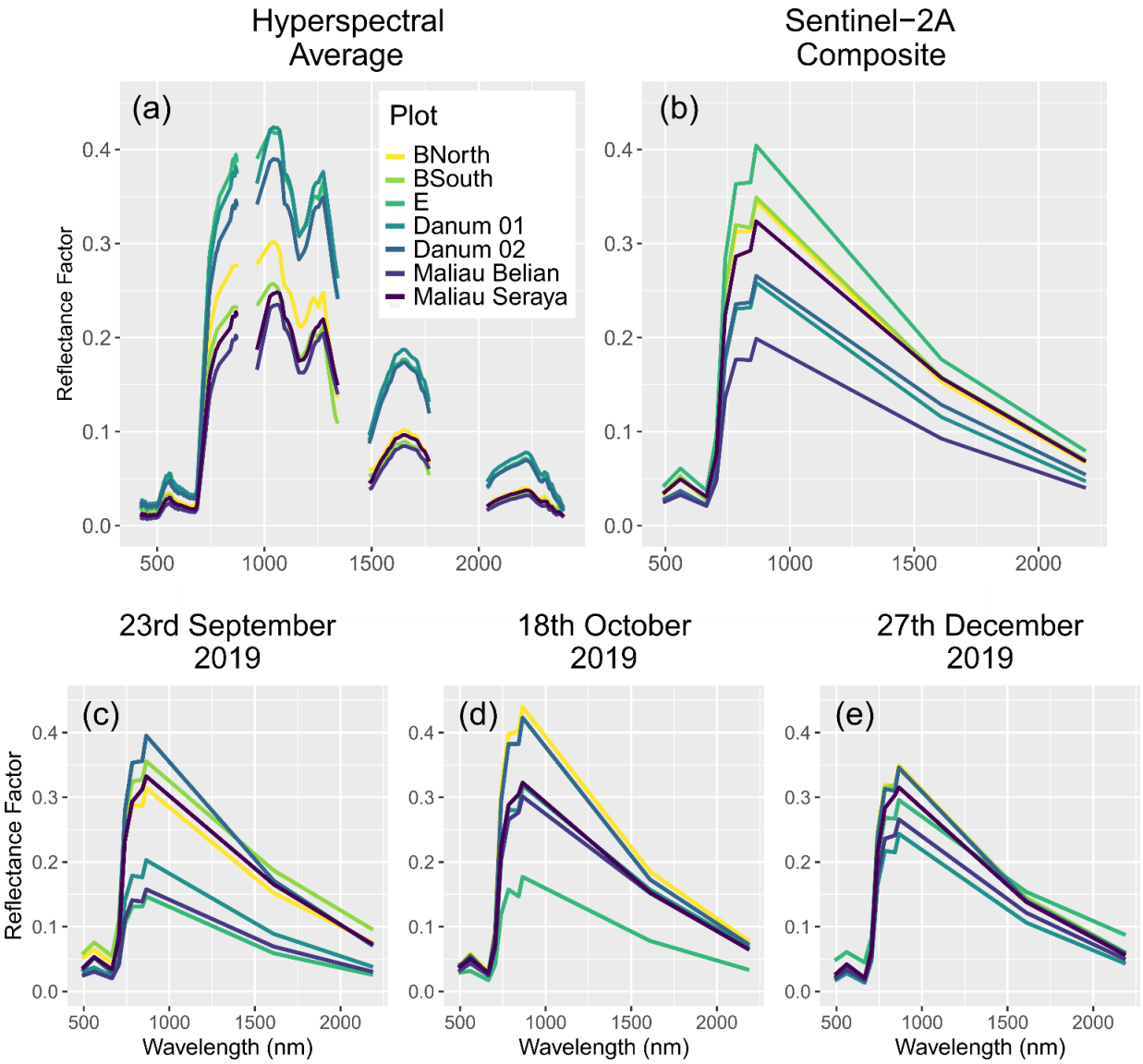


Figure 7.4: Average reflectance over seven tropical forest GEM plots computed using (a) airborne hyperspectral reflectance, (b) 10-band Sentinel-2 cloud-free composite, (c-e) 10-band Sentinel-2 cloud-free single scenes taken on 23rd September, 18th October and 27th December respectively. Clouds and cloud shadow removed from Sentinel imagery using Sen2Cor's Scene Classification Layer (SCL).

Evaluating spectral disagreements

Hyperspectral illumination conditions

A manual assessment of Sentinel imagery over the plots during October and November from 2015-2019 showed that clear vegetation was only present ~30% of the time (**Figure 7.5(a)**). Plots were more frequently covered in thick cloud (43%) and occasionally thin cloud (15%) and cloud shadow (12.5%).

Sen2Cor SCL

Evaluating the accuracy of Sentinel's cloud detection, we found that that the SCL tends to overestimate clear vegetation by around 50% (**Figure 7.5(b)**). On average, the plots exhibited clear vegetation ~40% of the time, yet the SCL classified them as displaying clear vegetation ~60% of the time. Undetected cloud cover was highest over the logged plots. From the confusion matrix (**Figure S7.3**), it can be seen that thin cloud and cloud shadow were most often inaccurately classified by the SCL as clear vegetation. However, no days of clear vegetation were missed by the SCL. Over the seven plots, the SCL had a total accuracy of 76%.

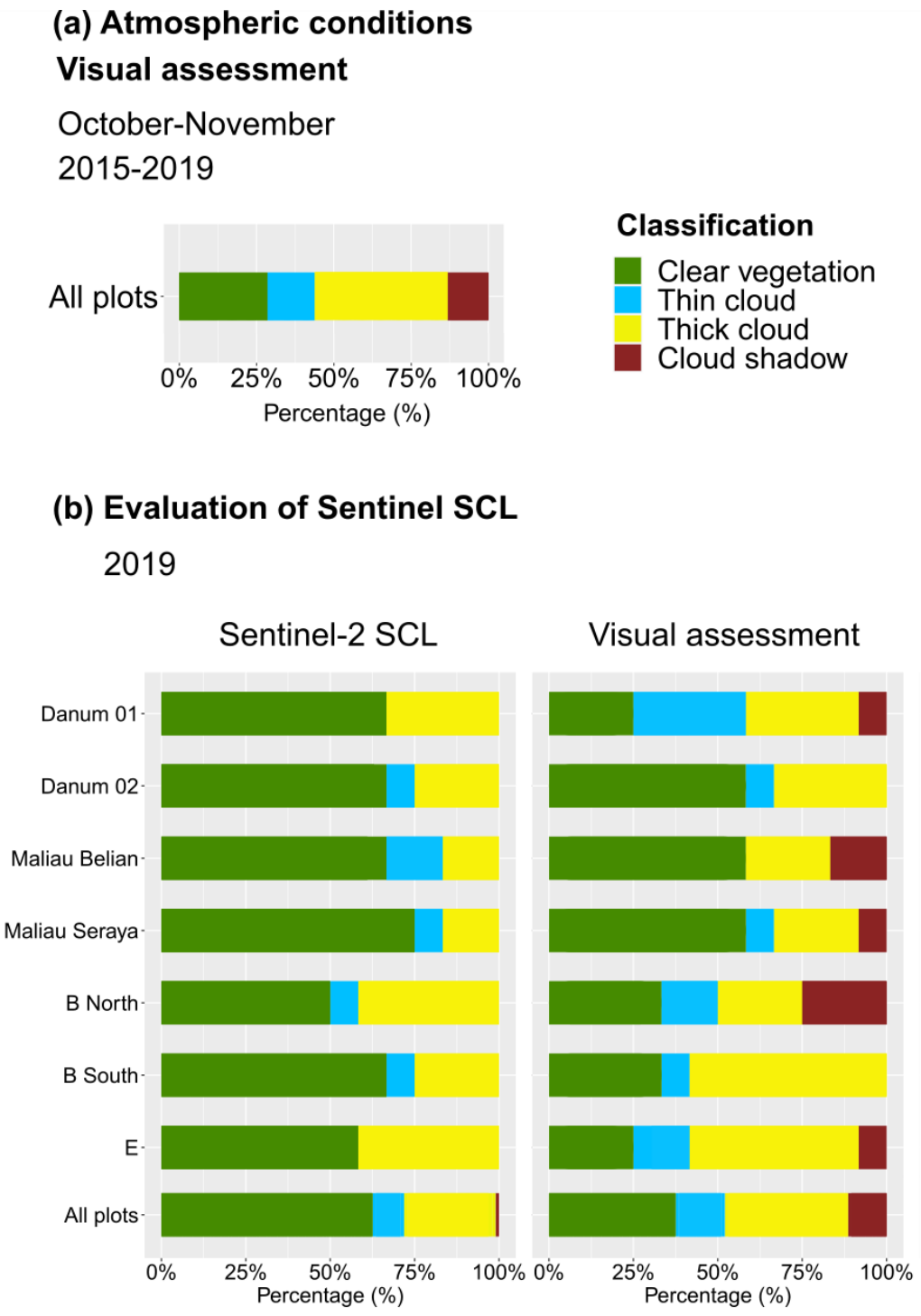


Figure 7.5*:* (a) Visual assessment of atmospheric conditions over seven tropical forest GEM plots for all Sentinel-2 imagery captured in October-November, 2015-2019. (b) Evaluation of Sen2Cor's cloud detection algorithm using its Scene Classification Layer (SCL), against a visual assessment for all Sentinel-2 images captured in 2019. See **Figure S7.3** for a confusion matrix of the results.

Sen2Cor surface reflectance values

Plotting a 5-day time-series data for Sentinel-2, reflectance values appear unstable and inconsistent in the humid tropical plots of Malaysia and Ankasa, with no clear seasonal consistency or signal (**Figure 7.6**). Unstable and inconsistent reflectance values is especially the case for the green and red bands, which show low values across all the Malaysian plots and large variations in reflectance over Ankasa 01. At these sites, the NIR band was more consistent, demonstrating a stronger distinction between the logged and unlogged plots in Malaysia, and between the wet evergreen, semi-deciduous and savannah forest in Ghana.

In contrast, a strong seasonal signal was apparent in the savannah plots of Brazil and Kogyae across all three bands. In Brazil, all plots showed increasing reflectance in the visible bands, with an inverse pattern in the NIR. In Ghana, a clear seasonal peak of high visible reflectance and low NIR was visible in March, likely related to water stress.

Visual impressions were confirmed by a statistical assessment of the data. Fitting each time-series line with a polynomial curve in order to distinguish surface reflectance error from an ecological signal, showed a large difference between the humid tropical forest and savannah plots (**Table 7.2**). The majority of the savannah plots exhibited a well-defined trend over the dry season (r^2>0.7), while most of the wet forest plots displayed high variation around a trend (r^2=0.04-0.66), or no trend whatsoever (r^2=0).

When calculating variance between nearby plots, nearly all paired plots tracked each other most closely in the NIR band. Tracking was especially pronounced for the humid tropical forest plots, further highlighting the inconstancy of visible Sentinel bands over humid tropical forest (**Table 7.3**).

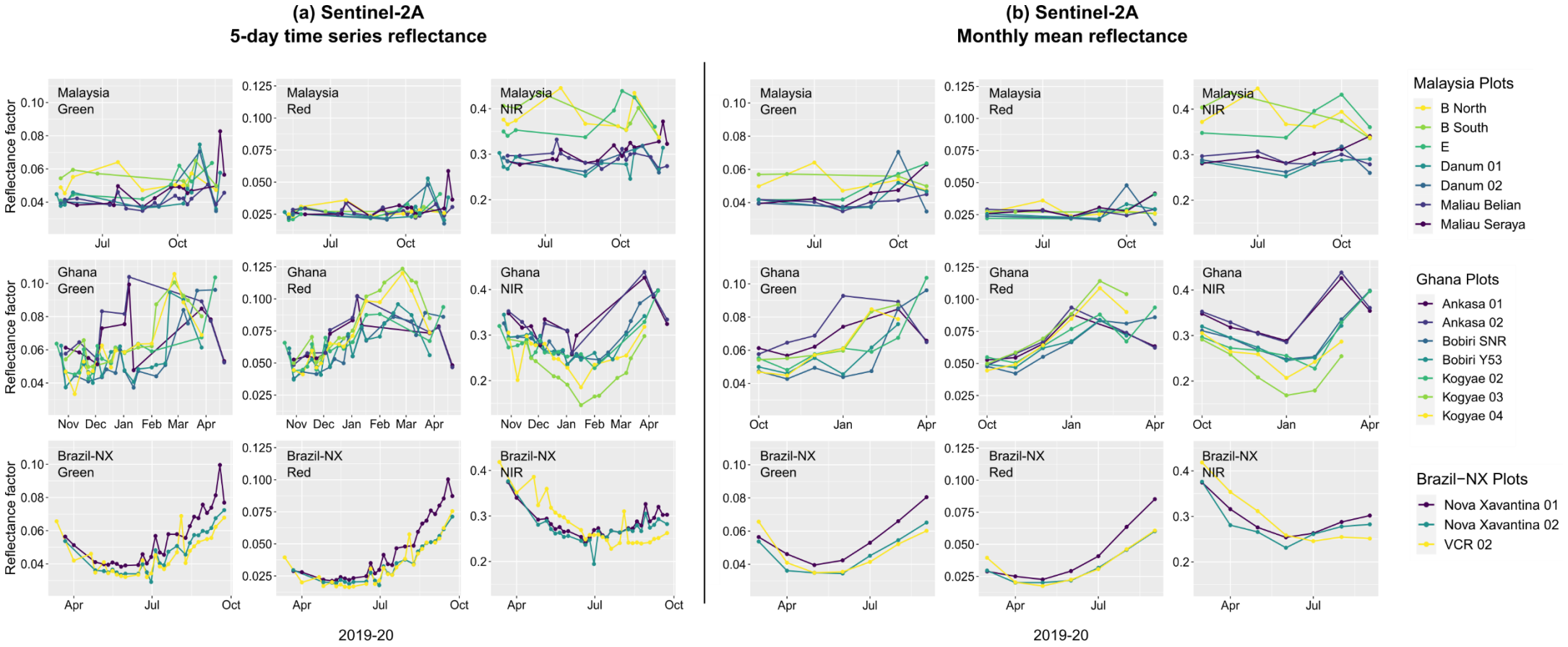


Figure 7.6: (a) Average 5-day time-series reflectance over 17 tropical forest plots in Malaysia, Ghana and Brazil. Reflectance plotted using Sentinel L2A imagery (Green band (Band 3), Red band (Band 4) and NIR band (Band 8 images)). (b) The same as (a) but for each plot, monthly reflectance values have been averaged. Images were visually checked to remove scenes exhibiting cloud cover and cloud shadow over plots.

Table 7.2*:* Relationship between 5-day Sentinel-2 reflectance from *Figure 7a* and a third-order polynomial curve of best-fit. Polynomial curve fitted in order to distinguish ecological signals from surface reflectance error. R-squared and RMSE values presented for each plot and each band (Green, Red and NIR). Entries are ranked according to R-squared and coloured according to forest type. Green=humid tropical forest, Yellow = semi-deciduous tropical forest and Orange = savannah-transition forest.

Site	Plot	Band	*n*	Adjusted R-squared	RMSE
Brazil	Nova Xavantina 01	Red	28	0.97	39
Brazil	Nova Xavantina 02	Red	26	0.96	29
Ghana	Kogyae 03	NIR	14	0.94	100
Brazil	Nova Xavantina 02	Green	26	0.92	32
Ghana	Kogyae 03	Red	14	0.91	65
Brazil	Nova Xavantina 01	Green	28	0.90	47
Ghana	Kogyae 04	Red	14	0.90	66
Ghana	Bobiri SNR	NIR	17	0.89	130
Brazil	VCR 02	Red	28	0.89	51
Brazil	Nova Xavantina 01	NIR	28	0.86	101
Ghana	Kogyae 02	NIR	16	0.85	134
Brazil	VCR 02	NIR	28	0.84	183
Ghana	Bobiri Y53	NIR	18	0.81	124
Ghana	Bobiri SNR	Red	17	0.78	72
Ghana	Bobiri SNR	Green	17	0.78	80
Ghana	Ankasa 03	Red	9	0.77	65
Brazil	Nova Xavantina 02	NIR	26	0.76	136
Malaysia	Danum 02	NIR	6	0.74	64
Brazil	VCR 02	Green	28	0.72	56
Ghana	Kogyae 03	Green	14	0.72	78
Ghana	Bobiri Y53	Red	18	0.69	83
Ghana	Ankasa 01	Red	11	0.69	74
Ghana	Kogyae 04	Green	14	0.66	91
Malaysia	E	Green	8	0.66	39
Ghana	Ankasa 03	Green	9	0.64	79
Malaysia	E	Red	8	0.60	31
Ghana	Kogyae 02	Green	16	0.60	79
Malaysia	Maliau Seraya	NIR	15	0.58	139
Ghana	Ankasa 03	NIR	9	0.58	259
Malaysia	E	NIR	8	0.55	187
Ghana	Ankasa 01	NIR	11	0.51	262
Ghana	Kogyae 02	Red	16	0.51	96
Malaysia	Maliau Seraya	Green	15	0.46	70
Malaysia	B South	NIR	7	0.36	182
Ghana	Kogyae 04	NIR	14	0.34	264
Malaysia	Maliau Seraya	Red	15	0.26	64
Ghana	Bobiri Y53	Green	18	0.20	122
Ghana	Ankasa 01	Green	11	0.07	125
Malaysia	B North	Red	9	0.06	29
Malaysia	Maliau Belian	Green	17	0.05	33
Malaysia	Danum 02	Green	6	0.04	76
Malaysia	Maliau Belian	NIR	17	0.00	143
Malaysia	Maliau Belian	Red	17	0.00	30
Malaysia	Danum 01	Green	12	0.00	95
Malaysia	Danum 01	NIR	12	0.00	206
Malaysia	B North	Green	9	0.00	47
Malaysia	Danum 01	Red	12	0.00	80
Malaysia	Danum 02	Red	6	0.00	69
Malaysia	B North	NIR	9	0.00	302
Malaysia	B South	Green	7	0.00	48

Table 7.3: Variance in Sentinel-2 reflectance from *Figure 7a* between plot pairs computed as a proportion of the overall variance for that GEM site (Brazil/Ghana/Malaysia). Results based on the premise that without random error, surface reflectance should vary less between two nearby plots than across all plots at a site. Proportion of variance results are ranked from smallest to largest and coloured by Sentinel-2 band (Green = Band 2, Red= Band 4 and Grey = Band 8). Results are only comparable between different bands of the same plot pair, not between different plots themselves, or between the same bands of different plot pairs.

Site	Plots	Proportion of variance	Band
Ghana	Ankasa 01+03	0.0028	NIR
Malaysia	Danum 01+02	0.0061	NIR
Malaysia	Danum 01+02	0.0119	Green
Malaysia	B North+South	0.0123	Red
Ghana	Ankasa 01+03	0.0163	Red
Malaysia	B North+South	0.0195	NIR
Malaysia	Maliau Belian+Seraya	0.0355	NIR
Malaysia	Danum 01+02	0.0465	Red
Ghana	Bobiri SNR+Y53	0.0479	NIR
Malaysia	B North+South	0.0504	Green
Ghana	Bobiri SNR+Y53	0.0554	Green
Ghana	Bobiri SNR+Y53	0.0557	Red
Brazil	Nova Xavantina 01+02	0.0655	NIR
Ghana	Kogyae 02+03	0.0686	Green
Ghana	Ankasa 01+03	0.1373	Green
Malaysia	Maliau Belian+Seraya	0.1468	Green
Malaysia	Maliau Belian+Seraya	0.3054	Red
Ghana	Kogyae 02+03	0.4084	NIR
Ghana	Kogyae 02+03	0.4985	Red
Brazil	Nova Xavantina 01+02	0.6677	Green
Brazil	Nova Xavantina 01+02	0.9778	Red

Predicting photosynthetic capacity

Tuning the hyperspectral imagery with the Sentinel composite, retained the detailed spectral information of the hyperspectral curves, but adjusted the magnitude of each spectrum to match the more credible spectral patterns of the Sentinel composite. (**Figure 7.7**).

The results of the PLSR analyses show that the tuned hyperspectral imagery produced the strongest spectra-trait models (also for A_{sat} and A_{max}, **Figure S7.4, Table S7.4**). $Jmax_{25}$ was estimated with the highest accuracy (r^2 = 0.43, RMSE = 20%), followed by $Vcmax_{25}$ (r^2 = 0.36, RMSE = 31%, **Figure 7.8, Table 7.4**). Increases in prediction accuracy were only significant for $Jmax_{25}$ (p <0.05), although using the tuned hyperspectral data produced spectral weighting for $Vcmax_{25}$ in the same spectral regions often used to predict biochemical traits associated with photosynthesis (e.g. chlorophyll and nitrogen concentrations, and LMA; Asner and Martin *et al*. 2008; Asner *et al*. 2015).

Estimating photosynthetic traits using the Sentinel imagery produced significantly lower results than for the hyperspectral imagery (r^2=0.28-0.32). Across all photosynthetic traits however, the Sentinel composite performed significantly better than the single Sentinel image (**Figure 7.8, Table 7.4**).

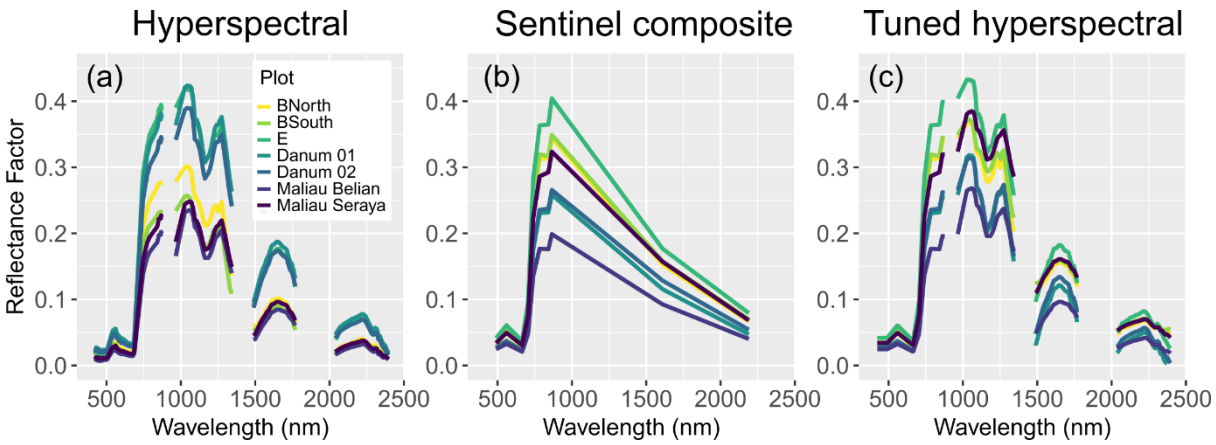


Figure 7.7: (a) Hyperspectral reflectance data over seven tropical forest plots reproduced from **Figure 7.4** (b) 10-band Sentinel-2 cloud-free composite over the same forest plot reproduced from **Figure 7.4**. (c) Tuned hyperspectral reflectance curves with magnitude based on Sentinel composite values.

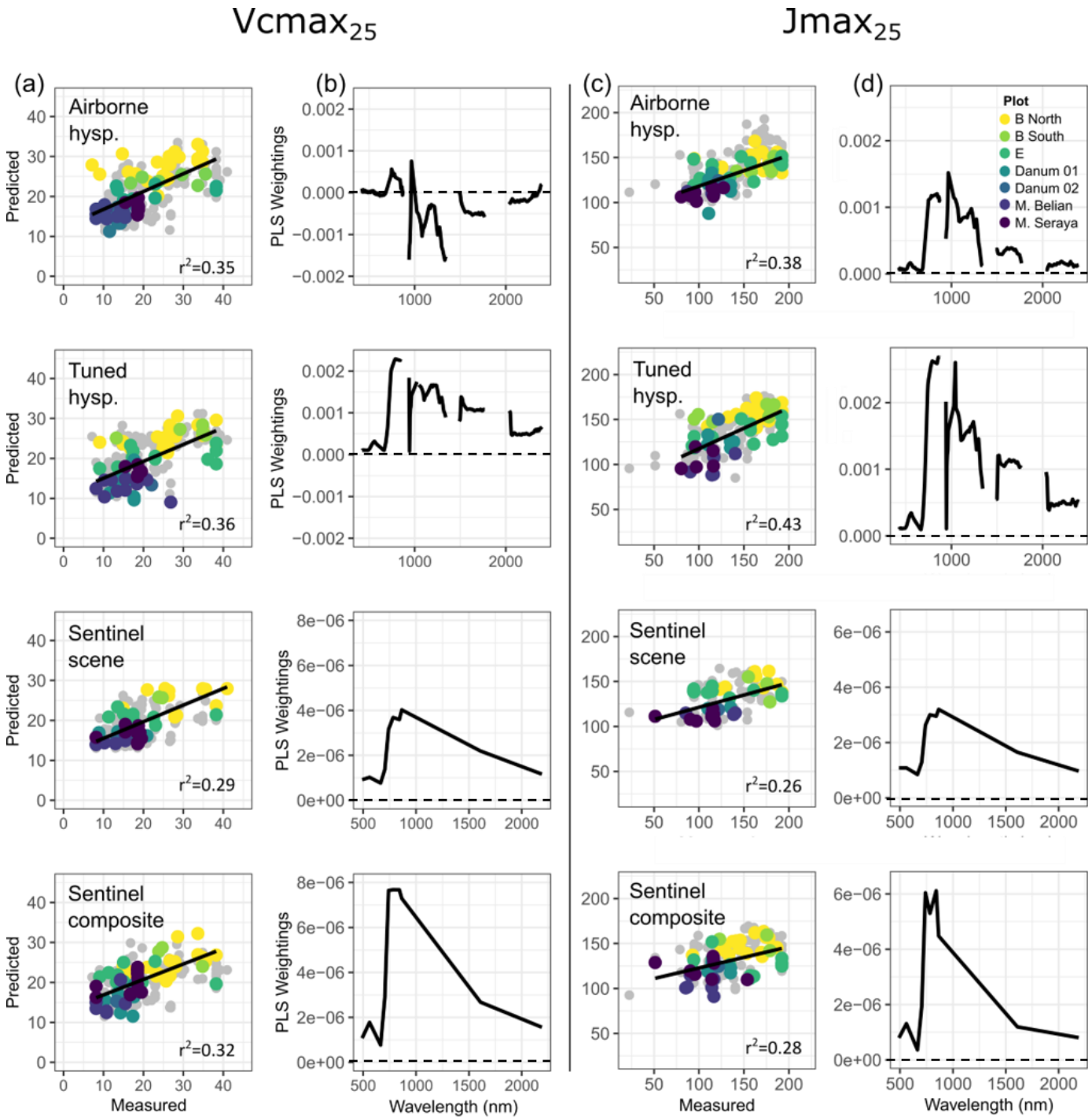


Figure 7.8*:* (a&c) Relationships between predicted versus measured plant photosynthetic capacity values ($Vcmax_{25}$ and $Jmax_{25}$) using PLSR analysis. (b&d): PLSR spectral weightings for each photosynthetic trait. Deviation from 0 indicates the parts of the spectrum that contribute most strongly to prediction. Traits were predicted four times using: (1) the airborne hyperspectral imagery (2) the hyperspectral imagery tuned with the Sentinel composite (**Figure 7.7**), (3) a single cloud-free Sentinel scene taken on 23rd of September, 2019 (4) a Sentinel time-series composite.

$Vcmax_{25}$	r^2	*%RMSE*	*RMSE*
Airborne hyperspectral	0.35	32	6.46
Tuned hyperspectral	0.36	31	6.38
Sentinel scene	0.29	34	6.41
Sentinel composite	0.32	34	6.78
$Jmax_{25}$			
Airborne hyperspectral	0.38	19	26.11
Tuned hyperspectral	0.43	20	26.78
Sentinel scene	0.26	22	29.03
Sentinel composite	0.28	22	29.85

Table 7.4*:* PLSR results of predicted *versus* measured plant photosynthesis values ($Vcmax_{25}$ and $Jmax_{25}$). Evaluated using r^2, RMSE as a percentage of the sample mean, and raw RMSE (μmol m^{-2} s^{-1}). Spectra were analysed four times using: (1) the airborne hyperspectral imagery (2) the hyperspectral imagery tuned with the Sentinel composite (**Figure 7.7**), (3) a single cloud-free Sentinel scene taken on 23rd of September, 2019 (4) a Sentinel time-series composite. All values are the average of 100 PLSR model runs.

7.7. Discussion

This study demonstrates the potential for using airborne hyperspectral imagery and Sentinel-2 data to map photosynthetic capacity ($Vcmax_{25}$ and $Jmax_{25}$) across a tropical forest. It also highlights some current limitations of optical remote sensing and presents best-practice recommendations for future airborne imaging studies.

Our results show that $Vcmax_{25}$ and $Jmax_{25}$ can be mapped using airborne hyperspectral imagery with a predictive performance of r^2=0.32 and 0.43 respectively (**Figure 7.8, Table 7.4**). Tuning the hyperspectral imagery to correct for spatially-inconsistent illumination conditions produced a significantly stronger spectra-trait model for $Jmax_{25}$ and spectral weighting for $Vcmax_{25}$ in the same spectral regions often used to predict biochemical traits associated with photosynthesis (e.g. chlorophyll and nitrogen concentrations, and LMA; Asner and Martin *et al*. 2008; Asner *et al*. 2015).

What are the most likely causes of cross-sensor disagreements between airborne hyperspectral imagery and Sentinel imagery taken in the same season?

Improved estimation of $Vcmax_{25}$ and $Jmax_{25}$ using the tuned hyperspectral imagery confirms that that a major reason behind the observed cross-sensor inconsistency is the sensitivity of the airborne hyperspectral imagery to varying illumination conditions. According to a 4-year visual assessment of the plots during the start of the wet season (when the hyperspectral imagery was collected), the plots were covered by cloud or cloud shadow around 70% of the time (**Figure 7.5(a)**) Cloud cover was not spatially consistent – there were few days when all the plots were clear or all the plots were cloudy – and even within a cloud cover classification, the type and depth of cloud cover varied, which results in differing amounts of radiation reaching the forest canopy (Marthews *et al*. 2012). Tropical forests occupy some of the cloudiest regions in the word with cloudless days the exception rather than the rule (Leigh 1999; Gabathuler *et al.* 2001); therefore, recording real-time illumination information alongside spectral values is vital to producing well-calibrated imagery that is comparable in both space and time.

In order to account for variable light conditions during a flight, most multispectral UAV cameras are supplied with a downwelling irradiance sensor, which is orientated upwards and synchronously records irradiance data along with imaging data (MicaSense 2017, 2019). During image processing, this metadata can be used to perform a scene-specific radiometric calibration to normalize images taken under different illumination conditions and calculate standardized reflectance values. In contrast, most hyperspectral imaging flights do not appear to follow such protocols. During post-processing, various methods have been employed to reduce or circumvent the issue of cloud cover. These include brightness normalization (dividing each image spectrum by its sum) to remove brightness differences while retaining absorption features (Janicke *et al*. 2019; Feilhauer *et al*. 2010), masking clouds manually (Draper *et al*. 2019) and using the shape of spectral curves rather than their amplitude to map volcanic material (Aufaristama *et al*. 2019). Whilst innovative, these approaches ultimately suffer from a lack of scene-specific metadata. Methods to reduce other sources of illumination variation, such as sun angle, BRDF effects, topographical effects and vegetation structure have been developed and are constantly improving (e.g Schläpfer *et al*. 2015; Singh *et al*. 2015; Weyermann *et al*. 2015; Brooke *et al*. 2018). We suggest that adding an irradiance sensor to an airborne hyperspectral imaging system would be one of the easiest and most effective measures to advance the accuracy and consistency of reflectance imagery. Hyperspectral imaging systems are still planned without such equipment (e.g. Sivertson *et al*. 2019). Scene-specific radiometric calibration should be an essential post-processing requirement for all future airborne spectroscopy studies.

What are the most likely causes of internal sensor disagreements between three single Sentinel images, and a Sentinel time-series composite, taken in the same season?

Across all the PLSR analyses, the Sentinel composite imagery estimated the photosynthetic parameters better than a single Sentinel scene. The Sentinel composite also displayed the most credible spectral patterns across the plots, such as spectral separability between the logged and old-growth forest. In comparison, the three single Sentinel scenes exhibited contrasting reflectance patterns to the Sentinel composite and

to each other (**Figure 7.4**). Evaluating the results of our Sen2Cor analyses we suggest this inconsistency is due to two main reasons.

Inaccuracy in Sentinel's cloud detection algorithm (SCL)

Firstly, the single Sentinel scenes were filtered using Sen2Cor's SCL but still contained unfiltered cloud or cloud shadowed pixels. According to our assessment, the SCL had an accuracy of 76% over the disturbance gradient (**Figure 7.5(b)**). Our results are similar to Sanchez *et al*. (2020), who compared 4 different cloud-detection algorithms and found Sen2Cor displayed the second highest accuracy in detecting cloud cover over the Amazon rainforest (79%) after Fmask4 (90%). Within an accuracy of 76% however, class-specific errors of omission and commission reduce the applied accuracy of the SCL. In a cloud-detection analysis, errors of clear-sky omission are less important than errors of commission, as omitting a few clear vegetation scenes will likely have less influence on overall results than including cloudy scenes. Yet, the SCL displays the opposite tendency. No clear vegetation scenes were omitted in our evaluation, yet 16 cloudy or cloud shadow scenes (31%) were incorrectly classified as clear vegetation (**Figure S7.3**). Therefore, a useful step to increase the practicality of the SCL for cloud-detection would be to make it more conservative in its classification of clear conditions, rather than conservative in its detection of clouds. In addition, when creating a Sentinel composite, it is important to use a median reflectance value, rather than a mean value, to further filter out cloudy and shadowed pixels.

Inaccurate surface reflectance values

The second reason why the three Sentinel scenes may display contrasting reflectance patterns is shown by a 5-day surface reflectance analysis, which found inconsistent surface reflectance values over humid tropical forests, even when all scenes had been manually assessed to exclude cloudy or cloud shadowed imagery (**Figure 7.6**). Unstable reflectance trends appeared to affect humid tropical forest more than savannah plots, and visible bands more than the NIR band, which was confirmed by a statistical analysis of the data (**Table 7.2, Table 7.3**). We hypothesise that these results are due to inaccurate Aerosol Optical Thickness (AOT) and Water Vapour (WV) retrieval by the Sen2Cor processor.

AOT is the measure of aerosols distributed within a column of air and WV is the measure of H_2O concentration within a column of air. Due to the scattering an absorbing effect of atmospheric aerosols and water vapour, accurate values for these parameters are required for deriving accurate surface reflectance from top of atmosphere (TOA) data. However, Li *et al.* (2018) found that for 40 AERONET sites over North America, Sen2Cor significantly overestimated AOT by 162%m with an overall correlation with *in situ* data of r=0.42. The validation of WV was better at r=0.94, although Sen2Cor underestimated the *in situ* data by 18%, with greater underestimation at higher values of WV (Main-Knorn *et al*. 2017). AOT and WV parameters are particularly important over tropical forests. Haze from biomass burning has increased rapidly in tropical regions during the last decades and is a particular problem over South-East Asia, linked to slash-and-burn activities and peat fires in Borneo, Kalimantan and Sumatra (Yanhon *et al*. 1996; Khan *et al*. 2020). In addition, wet tropical forests are one of the most humid regions of the world, with evaporation and transpiration processes playing a key role in global moisture recycling (Nobre *et al*. 1991; Kumagai *et al*. 2013). AOT and WV have also been found to interact, with high humidity increasing the hygroscopic composition of aerosols and leading to greater light scattering (Friendlander, 2000). These holistic effects may explain why Sen2Cor surface reflectance values are significantly less accurate over humid tropical forests than savannahs, with low signal-to-noise ratio confusing or entirely obscuring seasonal reflectance trends. In addition, shorter wavelengths of light are more strongly scattered by atmospheric aerosols and water vapour, explaining why reflectance patterns display more consistency in the NIR band. Ultimately, improved AOT and WV retrieval algorithms are required to increase surface reflectance accuracy over tropical forests. In the meantime, averaged monthly means may improve signal-to-noise ratio, but caution is still advised when using visible bands over tropical forests.

Which type of aerial imagery produces the strongest PLSR model for estimating photosynthetic capacity across a forest disturbance gradience in Malaysian Borneo?

All the airborne hyperspectral imagery displayed stronger relationships with $Vcmax_{25}$ and $Jmax_{25}$ (r^2= 0.35-0.43) than the Sentinel imagery (r^2= 0.21-0.32; **Figure 7.8, Table 7.4**). Photosynthetic capacity was correlated with hyperspectral reflectance at wavelengths broadly distributed across the full spectrum (i.e. 450—2500 nm), thus the

reduction in spectral information may explain the lower performance of Sentinel imagery. Similar spectral weightings were displayed by the hyperspectral and Sentinel analyses, with strong loadings observed at the red-edge (750nm) and across the NIR (100-1400nm), as well as near protein absorption features (910, 1020 and 1500nm), and light-absorbing pigments in the visible spectrum (640–670 nm). As leaf nitrogen primarily occurs in mesophyll proteins and chlorophylls with 70% of leaf nitrogen invested in compounds that support carbon fixation, these spectral weightings support a well-established link between leaf nutrient status and photosynthetic capacity (Field and Mooney, 1986; Evans, 1989; Reich *et al*. 1997; Dechant *et al*. 2017).

Does combining the spectral and spatial resolution of the aerial hyperspectral data with the stronger signal-to-noise ratio of the Sentinel time-series composite improve estimation of $Vcmax_{25}$ and $Jmax_{25}$?

Tuning the hyperspectral imagery to correct for spatially-inconsistent illumination conditions produced a significantly stronger spectra-trait models for $Jmax_{25}$, and more plausible spectral weightings for $Vcmax_{25}$ (**Figure 7.8, Table 7.4**). Nonetheless, the spectra-photosynthesis relationships found in our study were weaker than those found at the leaf level (e.g Serbin *et al.* 2012; Ainsworth *et al*. 2014; Heckmann *et al*. 2017), and lower than Serbin *et al*. (2015) found for 9 crop species using airborne hyperspectral imagery. Compared to leaf-level studies, our lower results are likely due to the effects of atmospheric and illumination conditions on spectral signatures and a higher number of species included in the analysis. Our results are more comparable to those of Doughty *et al*. (2011) who established leaf spectra-photosynthesis relationships with 11 tropical tree species at r^2=0.52. Lower results compared to Serbin *et al*. (2015) are likely due to a species rich and dense forest canopy that combines the spectral signatures of overlapping and adjacent tree crowns, compared to single-species crop fields.

However, our results were significantly stronger than Swinfield *et al.* (2019) reported for the same imagery. Over the same plots, Swinfield *et al*. (2019) developed relationships of r^2=0.08-0.20 for nitrogen and phosphorus concentrations and no relationship with SLA, despite leaf spectroscopy studies showing high correlation with these traits (e.g Asner and Martin *et al*. 2008; Serbin *et al*. 2014; Asner *et al*. 2015; Doughty *et al*. 2017).

Compared to Swinfield *et al*. (2019), our stronger spectra-trait models are probably due to high inter-plot variation in photosynthetic capacity (**Figure 7.3**). Swinfield *et al*. (2019) hypothesised that their trait predictions reflected spectral differences at the plot scale, rather than the tree crown level, hence why SLA, with low inter-plot variability, failed to display correlations with canopy spectra. As our photosynthetic traits displayed large differences between old growth and logged plots, these differences may explain our significantly stronger PLSR models for all traits.

7.8. Conclusion

Our results are promising for the large-scale mapping of photosynthetic capacity over tropical forests. Spatially-continuous datasets allow for the detection of drivers and patterns of trait variation at the landscape scale that would go undetected at the plot-level. Our airborne hyperspectral trait models captured 36-43% of the variation in photosynthetic capacity across a tropical disturbance gradient, and our Sentinel trait models captured 28-32% of variation in photosynthetic capacity across a tropical disturbance gradient. Our results are an improvement on TBMs, that typically assign one or two $Vcmax_{25}$ and $Jmax_{25}$ values to the entire tropical forest biome.

Reducing imagery quality issues associated with illumination and atmospheric effects will strengthen these results. Future airborne hyperspectral imagery studies would particularly benefit from improved radiometric calibration to account for varying illumination conditions during flights. Sentinel imagery would benefit from refined cloud-masking techniques and enhanced surface reflectance retrieval, especially over humid tropical forests.

7. 9. References

Aguirre-Gutiérrez, J., Rifai, S., Shenkin, A., Oliveras, I., Bentley, L.P., Svátek, M., Girardin, C.A.J., Both, S., Riutta, T. and Berenguer, E. (2020) Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data. *Remote Sensing of Environment,* 252, pp. 112122.

Ainsworth, E.A., Serbin, S.P., Skoneczka, J.A. and Townsend, P.A. (2014) Using leaf optical properties to detect ozone effects on foliar biochemistry. *Photosynthesis Research,* 119 (1), pp. 65-76.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn, N.R. and Llactayo, W. (2017) Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science,* 355 (6323), pp. 385-389.

Asner, G.P. and Martin, R.E. (2015) Spectroscopic remote sensing of non-structural carbohydrates in forest canopies. *Remote Sensing,* 7 (4), pp. 3526-3547.

Asner, G.P., Martin, R.E., Tupayachi, R., Anderson, C.B., Sinca, F., Carranza-Jiménez, L. and Martinez, P. (2014) Amazonian functional diversity from forest canopy chemical assembly. *Proceedings of the National Academy of Sciences,* 111 (15), pp. 5604-5609.

Aufaristama, M., Hoskuldsson, A., Ulfarsson, M.O., Jonsdottir, I. and Thordarson, T. (2019) The 2014–2015 Lava Flow Field at Holuhraun, Iceland: Using Airborne Hyperspectral Remote Sensing for Discriminating the Lava Surface. *Remote Sensing,* 11 (5), pp. 476.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M.A., Baldocchi, D. and Bonan, G.B. (2010) Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science,* 329 (5993), pp. 834-838.

Bonan, G.B. (2008) Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science,* 320 (5882), pp. 1444-1449.

Both, S., Riutta, T., Paine, C.T., Elias, D.M.O., Cruz, R.S., Jain, A., Johnson, D., Kritzler, U.H., Kuntz, M. and Majalap-Lee, N. (2019) Logging and soil nutrients independently explain plant trait expression in tropical forests. *New Phytologist,* 221 (4), pp. 1853-1865.

Brook, A., Polinova, M. and Ben-Dor, E. (2018) Fine tuning of the SVC method for airborne hyperspectral sensors: The BRDF correction of the calibration nets targets. *Remote Sensing of Environment,* 204, pp. 861-871.

Cheng, Y., Vrieling, A., Fava, F., Meroni, M., Marshall, M. and Gachoki, S. (2020) Phenology of short vegetation cycles in a Kenyan rangeland from PlanetScope and Sentinel-2. *Remote Sensing of Environment,* 248, pp. 112004.

Chuvieco, E., Pettinari, M.L., Koutsias, N., Forkel, M., Hantson, S. and Turco, M. (2021) Human and climate drivers of global biomass burning variability. *Science of The Total Environment,* 779, pp. 146361.

Curran, P.J. (1989) Remote sensing of foliar chemistry. *Remote Sensing of Environment,* 30 (3), pp. 271-278.

De Kauwe, M.G., Lin, Y., Wright, I.J., Medlyn, B.E., Crous, K.Y., Ellsworth, D.S., Maire, V., Prentice, I.C., Atkin, O.K. and Rogers, A. (2016) A test of the 'one-point method'for estimating maximum carboxylation capacity from field-measured, light-saturated photosynthesis. *New Phytologist,* 210 (3), pp. 1130-1144.

Dechant, B., Cuntz, M., Vohland, M., Schulz, E. and Doktor, D. (2017) Estimation of photosynthesis traits from leaf reflectance spectra: Correlation to nitrogen content as the dominant mechanism. *Remote Sensing of Environment,* 196, pp. 279-292.
Doughty, C.E., Asner, G.P. and Martin, R.E. (2011) Predicting tropical plant physiology from leaf and canopy spectroscopy. *Oecologia,* 165 (2), pp. 289-299.
Doughty, C.E., Santos-Andrade, P.E., Goldsmith, G.R., Blonder, B., Shenkin, A., Bentley, L.P., Chavana-Bryant, C., Huaraca-Huasco, W., Díaz, S. and Salinas, N. (2017) Can leaf spectroscopy predict leaf and forest traits along a Peruvian tropical forest elevation gradient? *Journal of Geophysical Research: Biogeosciences,* 122 (11), pp. 2952-2965.
Draper, F.C., Baraloto, C., Brodrick, P.G., Phillips, O.L., Martinez, R.V., Honorio Coronado, E.N., Baker, T.R., Zárate Gómez, R., Amasifuen Guerra, C.A. and Flores, M. (2019) Imaging spectroscopy predicts variable distance decay across contrasting Amazonian tree communities. *Journal of Ecology,* 107 (2), pp. 696-710.
Duursma, R.A. (2015) Plantecophys-an R package for analysing and modelling leaf gas exchange data. *PloS one,* 10 (11), pp. e0143346.
Eck, T.F., Holben, B.N., Reid, J.S., Mukelabai, M.M., Piketh, S.J., Torres, O., Jethva, H.T., Hyer, E.J., Ward, D.E. and Dubovik, O. (2013) A seasonal trend of single scattering albedo in southern African biomass-burning particles: Implications for satellite products and estimates of emissions for the world's largest biomass-burning source. *Journal of Geophysical Research: Atmospheres,* 118 (12), pp. 6414-6432.
Evans, J.R. (1989) Photosynthesis and nitrogen relationships in leaves of C 3 plants. *Oecologia,* 78 (1), pp. 9-19.
Farquhar, G.D., von Caemmerer, S.v. and Berry, J.A. (1980) A biochemical model of photosynthetic CO 2 assimilation in leaves of C 3 species. *Planta,* 149 (1), pp. 78-90.
Feilhauer, H., Asner, G.P., Martin, R.E. and Schmidtlein, S. (2010) Brightness-normalized partial least squares regression for hyperspectral data. *Journal of Quantitative Spectroscopy and Radiative Transfer,* 111 (12-13), pp. 1947-1957.
Field, C.H. and Mooney, H.A. (1986) Photosynthesis--nitrogen relationship in wild plants. Cambridge Cambridgeshire]: Cambridge University Press, c1986.
Friendlander, S.K. (2000) Smoke, dust and haze: Fundamentals of aerosol dynamics.
Gabathuler, M., Marty, C.A. and Hanselmann, K.W. (2001) Parameterization of incoming longwave radiation in high-mountain environments. *Physical Geography,* 22 (2), pp. 99-114.
Gara, T.W., Darvishzadeh, R., Skidmore, A.K., Wang, T. and Heurich, M. (2019) Accurate modelling of canopy traits from seasonal Sentinel-2 imagery based on the vertical distribution of leaf traits. *ISPRS journal of photogrammetry and remote sensing,* 157, pp. 108-123.
Guan, K., Berry, J.A., Zhang, Y., Joiner, J., Guanter, L., Badgley, G. and Lobell, D.B. (2016) Improving the monitoring of crop productivity using spaceborne solar-induced fluorescence. *Global Change Biology,* 22 (2), pp. 716-726.
Guanter, L., Alonso, L., Gómez-Chova, L., Amorós-López, J., Vila, J. and Moreno, J. (2007) Estimation of solar-induced vegetation fluorescence from space measurements. *Geophysical Research Letters,* 34 (8).
Gvozdevaite, A., Oliveras, I., Domingues, T.F., Peprah, T., Boakye, M., Afriyie, L., da Silva Peixoto, K., de Farias, J., Almeida de Oliveira, E. and Almeida Farias, C.C. (2018)

Leaf-level photosynthetic capacity dynamics in relation to soil and foliar nutrients along forest–savanna boundaries in Ghana and Brazil. *Tree physiology,* 38 (12), pp. 1912-1925.
Haaland, D.M. and Thomas, E.V. (1988) Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry,* 60 (11), pp. 1193-1202.
Heckmann, D., Schlüter, U. and Weber, A.P.M. (2017) Machine learning techniques for predicting crop photosynthetic capacity from leaf reflectance spectra. *Molecular plant,* 10 (6), pp. 878-890.
Hédl, R., Svátek, M., Dančák, M., Rodzay, A.W., Salleh AB, M. and Kamariah, A.S. (2009) A new technique for inventory of permanent plots in tropical forests: a case study from lowland dipterocarp forest in Kuala Belalong, Brunei Darussalam. *Blumea-Biodiversity, Evolution and Biogeography of Plants,* 54 (1-2), pp. 124-130.
Jänicke, C., Okujeni, A., Cooper, S., Clark, M., Hostert, P. and van der Linden, S. (2020) Brightness gradient-corrected hyperspectral image mosaics for fractional vegetation cover mapping in northern California. *Remote Sensing Letters,* 11 (1), pp. 1-10.
Khan, M.F., Hamid, A.H., Ab Rahim, H., Maulud, K.N.A., Latif, M.T., Nadzir, M.S.M., Sahani, M., Qin, K., Kumar, P. and Varkkey, H. (2020) El Niño driven haze over the Southern Malaysian Peninsula and Borneo. *Science of The Total Environment,* 730, pp. 139091.
Kokaly, R.F. (2001) Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment,* 75 (2), pp. 153-161.
Kumagai, T., Kanamori, H. and Yasunari, T. (2013) Deforestation-induced reduction in rainfall. *Hydrological Processes,* 27 (25), pp. 3811-3814.
Lee, C.M., Cable, M.L., Hook, S.J., Green, R.O., Ustin, S.L., Mandl, D.J. and Middleton, E.M. (2015) An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. *Remote Sensing of Environment,* 167, pp. 6-19.
Leigh, E.G. (1999) *Tropical forest ecology: a view from Barro Colorado Island.* Oxford University Press on Demand.
Li, Y., Chen, J., Ma, Q., Zhang, H.K. and Liu, J. (2018) Evaluation of Sentinel-2A surface reflectance derived using Sen2Cor in North America. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11 (6), pp. 1997-2021.
Ma, X., Mahecha, M.D., Migliavacca, M., van der Plas, F., Benavides, R., Ratcliffe, S., Kattge, J., Richter, R., Musavi, T. and Baeten, L. (2019) Inferring plant functional diversity from space: the potential of Sentinel-2. *Remote Sensing of Environment,* 233, pp. 111368.
MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F. and Ciais, P. (2015) Using satellite data to improve the leaf phenology of a global terrestrial biosphere model. *Biogeosciences,* 12 (23), pp. 7185-7208.
Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U. and Gascon, F. (2017) Sen2Cor for sentinel-2. International Society for Optics and Photonics, pp. 1042704.
Malhi, Y., Girardin, C., Metcalfe, D.B., Doughty, C.E., Aragão, L.E., Rifai, S.W., Oliveras, I., Shenkin, A., Aguirre-Gutiérrez, J. and Dahlsjö, C.A. (2021) The Global Ecosystems Monitoring network: Monitoring ecosystem productivity and carbon cycling across the tropics. *Biological Conservation,* 253, pp. 108889.

Malhi, Y. (2012) The productivity, metabolism and carbon cycle of tropical forest vegetation. *Journal of Ecology,* 100 (1), pp. 65-75.
Marsh, C.W. and Greer, A.G. (1992) Forest land-use in Sabah, Malaysia: an introduction to Danum Valley. *Philosophical Transactions of the Royal Society of London.Series B: Biological Sciences,* 335 (1275), pp. 331-339.
Marthews, T.R., Malhi, Y. and Iwata, H. (2012) Calculating downward longwave radiation under clear and cloudy conditions over a tropical lowland forest site: an evaluation of model schemes for hourly data. *Theoretical and applied climatology,* 107 (3), pp. 461-477.
Meacham-Hensold, K., Montes, C.M., Wu, J., Guan, K., Fu, P., Ainsworth, E.A., Pederson, T., Moore, C.E., Brown, K.L. and Raines, C. (2019) High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. *Remote Sensing of Environment,* 231, pp. 111176.
Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R. and Moreno, J. (2009) Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. *Remote Sensing of Environment,* 113 (10), pp. 2037-2051.
MicaSense (2017) https://support.micasense.com/hc/en-us/articles/115002782008-Light-Sensors-the-Basics-DLS-and-Sunshine-Sensor-. Accessed 04/12/20.
MicaSense (2019) https://support.micasense.com/hc/en-us/articles/219901327-What-does-RedEdge-s-Downwelling-Light-Sensor-do-for-my-data-. Accessed 04/12/20.
Nainar, A., Bidin, K., Walsh, R.P.D., Ewers, R.M. and Reynolds, G. (2015) Variations in suspended sediment yield and dynamics in catchments of differing land-use in Sabah. *Transactions on Science and Technology,* 2 (1), pp. 1-19.
Nobre, C.A., Sellers, P.J. and Shukla, J. (1991) Amazonian deforestation and regional climate change. *Journal of Climate,* 4 (10), pp. 957-988.
Oliveras, I., Bentley, L., Fyllas, N.M., Gvozdevaite, A., Shenkin, A.F., Peprah, T., Morandi, P., Peixoto, K.S., Boakye, M. and Adu-Bredu, S. (2020) The influence of taxonomy and environment on leaf trait variation along tropical abiotic gradients. *Frontiers in Forests and Global Change,* 3, pp. 18.
Reich, P.B., Walters, M.B. and Ellsworth, D.S. (1997) From tropics to tundra: global convergence in plant functioning. *Proceedings of the National Academy of Sciences,* 94 (25), pp. 13730-13734.
Rifai, S.W., Li, S. and Malhi, Y. (2019) Coupling of El Niño events and long-term warming leads to pervasive climate extremes in the terrestrial tropics. *Environmental Research Letters,* 14 (10), pp. 105002.
Riutta, T., Malhi, Y., Kho, L.K., Marthews, T.R., Huaraca Huasco, W., Khoo, M., Tan, S., Turner, E., Reynolds, G. and Both, S. (2018) Logging disturbance shifts net primary productivity and its allocation in Bornean tropical forests. *Global Change Biology,* 24 (7), pp. 2913-2928.
Rogers, A. (2014) The use and misuse of V c, max in Earth System Models. *Photosynthesis Research,* 119 (1), pp. 15-29.
Sanchez, A.H., Picoli, M.C.A., Camara, G., Andrade, P.R., Chaves, M.E.D., Lechler, S., Soares, A.R., Marujo, R.F.B., Simões, R.E.O. and Ferreira, K.R. (2020) Comparison of Cloud cover detection algorithms on sentinel–2 images of the amazon tropical forest. *Remote Sensing,* 12 (8), pp. 1284.

Schläpfer, D., Richter, R. and Feingersh, T. (2014) Operational BRDF effects correction for wide-field-of-view optical scanners (BREFCOR). *IEEE Transactions on Geoscience and Remote Sensing,* 53 (4), pp. 1855-1864.
Serbin, S.P., Dillaway, D.N., Kruger, E.L. and Townsend, P.A. (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *Journal of experimental botany,* 63 (1), pp. 489-502.
Serbin, S.P., Singh, A., Desai, A.R., Dubois, S.G., Jablonski, A.D., Kingdon, C.C., Kruger, E.L. and Townsend, P.A. (2015) Remotely estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using imaging spectroscopy. *Remote Sensing of Environment,* 167, pp. 78-87.
Serbin, S.P., Singh, A., McNeil, B.E., Kingdon, C.C. and Townsend, P.A. (2014) Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecological Applications,* 24 (7), pp. 1651-1669.
Singh, A., Serbin, S.P., McNeil, B.E., Kingdon, C.C. and Townsend, P.A. (2015) Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications,* 25 (8), pp. 2180-2197.
Streets, D.G., Yan, F., Chin, M., Diehl, T., Mahowald, N., Schultz, M., Wild, M., Wu, Y. and Yu, C. (2009) Anthropogenic and natural contributions to regional trends in aerosol optical depth, 1980–2006. *Journal of Geophysical Research: Atmospheres,* 114.
Swinfield, T., Both, S., Riutta, T., Bongalov, B., Elias, D., Majalap-Lee, N., Ostle, N., Svátek, M., Kvasnica, J. and Milodowski, D. (2020) Imaging spectroscopy reveals the effects of topography and logging on the leaf chemistry of tropical forest canopy trees. *Global Change Biology,* 26 (2), pp. 989-1002.
Thomson, E.R., Malhi, Y., Bartholomeus, H., Oliveras, I., Gvozdevaite, A., Peprah, T., Suomalainen, J., Quansah, J., Seidu, J. and Adonteng, C. (2018) Mapping the Leaf Economic Spectrum across West African Tropical Forests Using UAV-Acquired Hyperspectral Imagery. *Remote Sensing,* 10 (10), pp. 1532.
Thomson, E.R., Spiegel, M.P., Althuizen, I.H.J., Bass, P., Chen, S., Chmurzynski, A., Halbritter, A.H., Henn, J.J., Jónsdóttir, I.S. and Klanderud, K. (2021) Multiscale mapping of plant functional groups and plant traits in the High Arctic using field spectroscopy, UAV imagery and Sentinel-2A data. *Environmental Research Letters,* 16 (5), pp. 055006.
Tobler, W.R. (1970) A computer movie simulating urban growth in the Detroit region. *Economic geography,* 46, pp. 234-240.
Ustin, S.L., Gitelson, A.A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J.A. and Zarco-Tejada, P. (2009) Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sensing of Environment,* 113, pp. S67-S77.
Walsh, R. and Newbery, D.M. (1999) The ecoclimatology of Danum, Sabah, in the context of the world's rainforest regions, with particular reference to dry periods and their impact. *Philosophical Transactions of the Royal Society of London.Series B: Biological Sciences,* 354 (1391), pp. 1869-1883.
Weyermann, J.ö., Kneubühler, M., Schläpfer, D. and Schaepman, M.E. (2015) Minimizing reflectance anisotropy effects in airborne spectroscopy data using Ross–Li model inversion with continuous field land cover stratification. *IEEE Transactions on Geoscience and Remote Sensing,* 53 (11), pp. 5814-5823.

Wu, J., Rogers, A., Albert, L.P., Ely, K., Prohaska, N., Wolfe, B.T., Oliveira Jr, R.C., Saleska, S.R. and Serbin, S.P. (2019) Leaf reflectance spectroscopy captures variation in carboxylation capacity across species, canopy environment and leaf age in lowland moist tropical forests. *New Phytologist,* 224 (2), pp. 663-674.

Wullschleger, S.D., Epstein, H.E., Box, E.O., Euskirchen, E.S., Goswami, S., Iversen, C.M., Kattge, J., Norby, R.J., Van Bodegom, P.M. and Xu, X. (2014) Plant functional types in Earth system models: past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. *Annals of botany,* 114 (1), pp. 1-16.

Yang, K., Ryu, Y., Dechant, B., Berry, J.A., Hwang, Y., Jiang, C., Kang, M., Kim, J., Kimm, H. and Kornfeld, A. (2018) Sun-induced chlorophyll fluorescence is more strongly related to absorbed light than to photosynthesis at half-hourly resolution in a rice paddy. *Remote Sensing of Environment,* 216, pp. 658-673.

Yang, L., Song, J., Han, L., Wang, X. and Wang, J. (2020) Reconstruction of High-Temporal-and High-Spatial-Resolution Reflectance Datasets Using Difference Construction and Bayesian Unmixing. *Remote Sensing,* 12 (23), pp. 3952.

Yanhong, T., Naoki, K., Akio, F. and Awang, M. (1996) Light reduction by regional haze and its effect on simulated leaf photosynthesis in a tropical forest of Malaysia. *Forest Ecology and Management,* 89 (1-3), pp. 205-211.

Zekoll, V., Main-Knorn, M., Louis, J., Frantz, D., Richter, R. and Pflug, B. (2021) Comparison of masking algorithms for sentinel-2 imagery. *Remote Sensing,* 13 (1), pp. 137.

Zaehle, S., Sitch, S., Smith, B. and Hatterman, F. (2005) Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics. *Global Biogeochemical Cycles,* 19 (3).

7.10. Supplementary material

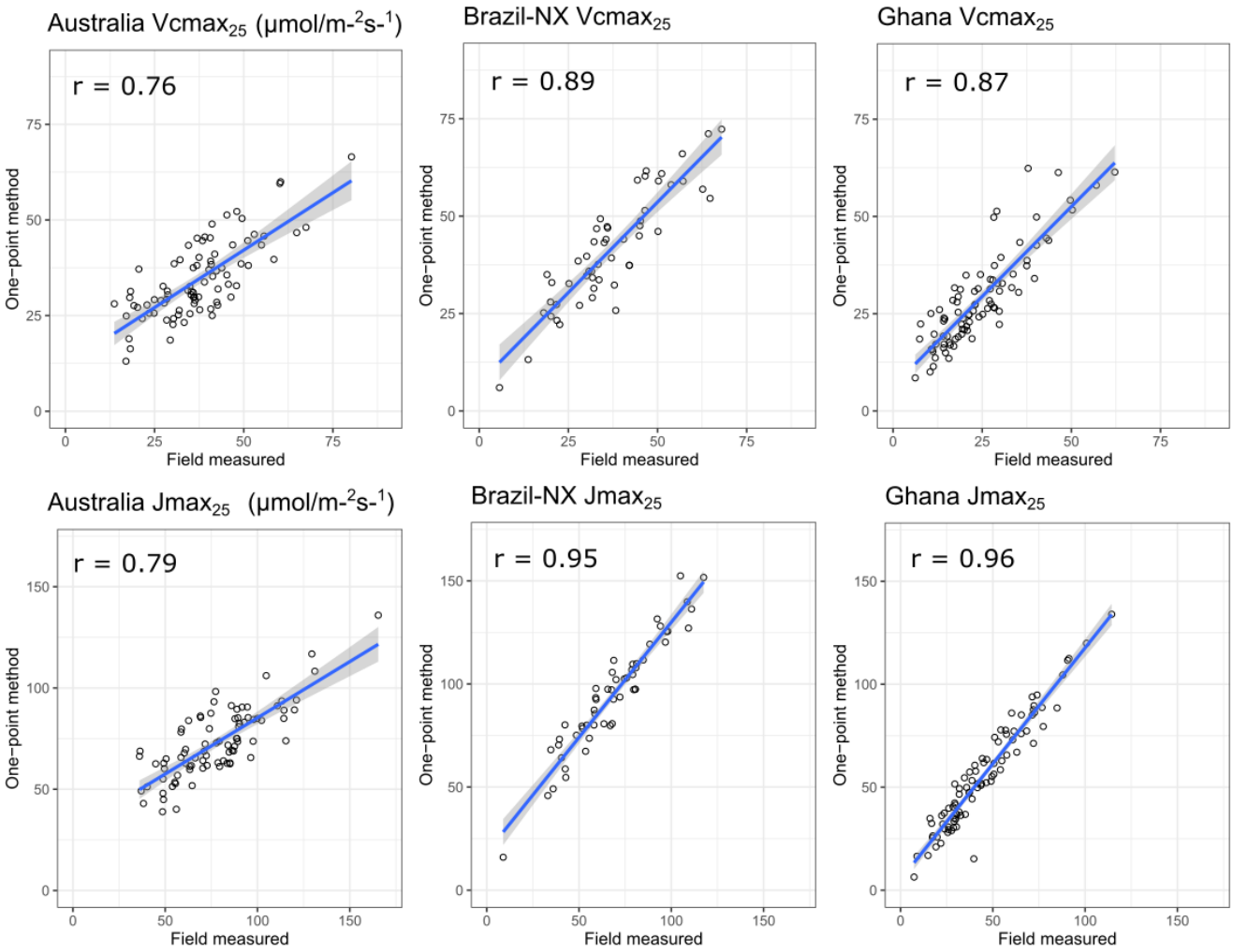


Figure S7.1*:* Prediction accuracy of de Kauwe *et al.*'s (2016) 'one-point' method compared against $Vcmax_{25}$ and $Jmax_{25}$ derived from field-measured A-Ci-curves.

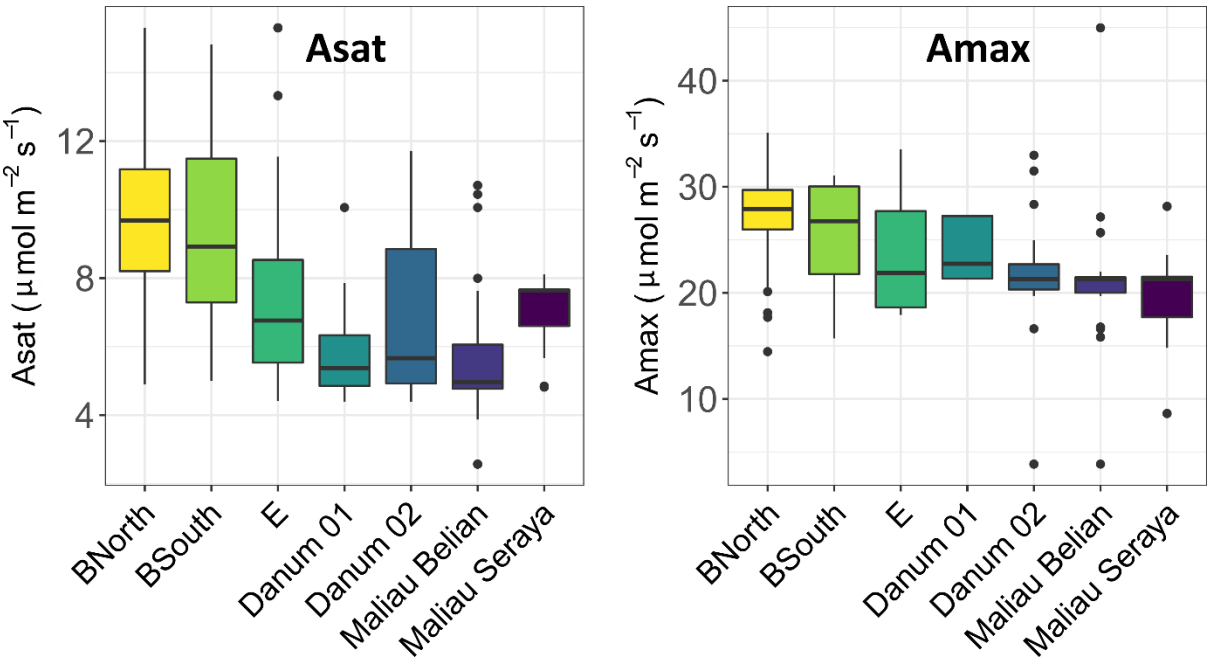


Figure S7.2: A_{sat} and A_{max} values across seven one-hectare GEM plots in Malaysian Borneo representing a forest disturbance gradient from logged to old-growth forest. Plots are ordered from heavily logged (BNorth and BSouth), to lightly logged (E), to old-growth forest (Danum and Maliau).

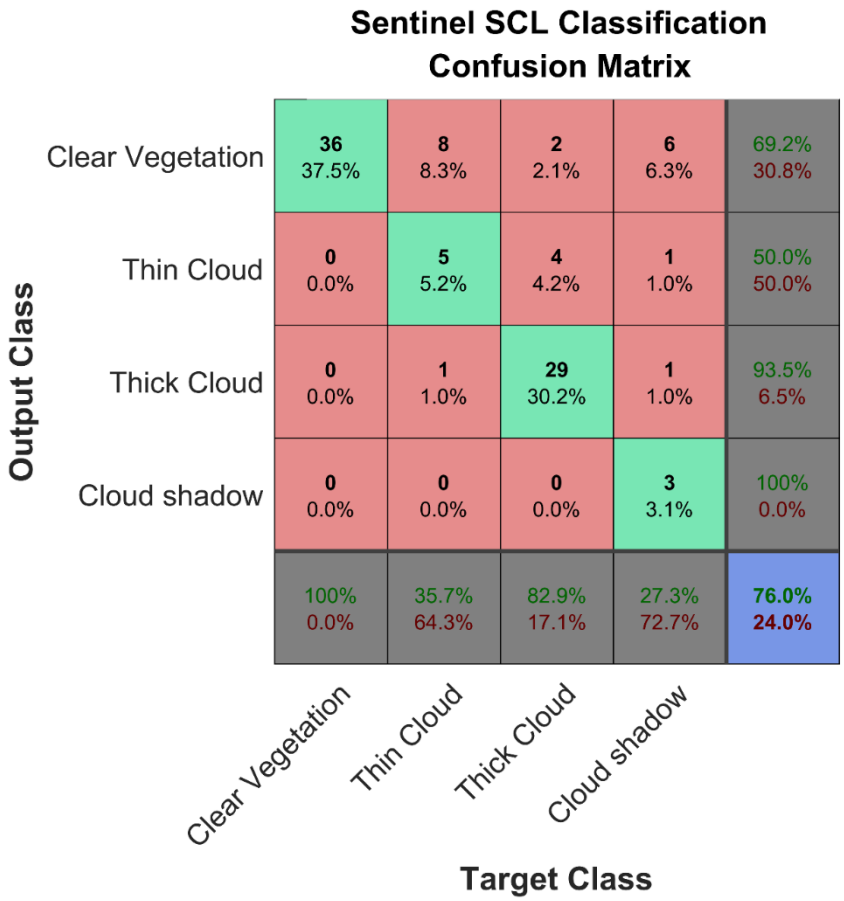


Figure S7.3*:* Confusion matrix for Sentinel SCL classification *vs* a visual assessment of atmospheric conditions over seven tropical forest plots in Malaysian Borneo in 2019.

Table S7.1*:* Relationship between 5-day Sentinel-2A reflectance from **Figure 7.5a** and a second-order polynomial curve of best-fit. Polynomial curve fitted in order to distinguish ecological signals from random error. r^2 and RMSE values presented for each plot and each band (Green, Red and NIR). Entries are ranked according to r^2 and coloured according to forest type. Green=humid tropical forest, Yellow = semi-deciduous tropical forest and Orange = savannah-transition forest.

Site	Plot	Band	*n*	Adjusted R-squared	RMSE
Brazil	Nova Xavantina 01	Red	28	0.97	39
Brazil	Nova Xavantina 02	Red	26	0.96	29
Brazil	Nova Xavantina 02	Green	26	0.91	35
Brazil	Nova Xavantina 01	Green	28	0.90	49
Brazil	VCR 02	Red	28	0.89	52
Ghana	Bobiri SNR	NIR	17	0.87	146
Brazil	VCR 02	NIR	28	0.84	186
Ghana	Kogyae 03	NIR	14	0.84	169
Brazil	Nova Xavantina 01	NIR	28	0.83	113
Ghana	Kogyae 02	NIR	16	0.80	160
Ghana	Bobiri SNR	Green	17	0.79	81
Ghana	Bobiri Y53	NIR	18	0.74	152
Ghana	Kogyae 03	Red	14	0.73	120
Ghana	Bobiri SNR	Red	17	0.73	83
Malaysia	E	Green	8	0.72	39
Ghana	Kogyae 04	Red	14	0.71	118
Ghana	Ankasa 03	Red	9	0.70	81
Brazil	VCR 02	Green	28	0.69	60
Brazil	Nova Xavantina 02	NIR	26	0.68	162
Ghana	Kogyae 03	Green	14	0.64	93
Ghana	Kogyae 02	Green	16	0.62	80
Malaysia	E	Red	8	0.62	34
Malaysia	Maliau Seraya	NIR	15	0.61	140
Ghana	Ankasa 01	Red	11	0.61	89
Ghana	Kogyae 04	Green	14	0.59	104
Malaysia	B South	NIR	7	0.50	186
Malaysia	Maliau Seraya	Green	15	0.48	72
Ghana	Kogyae 02	Red	16	0.46	104
Ghana	Ankasa 03	Green	9	0.45	106
Ghana	Bobiri Y53	Red	18	0.45	114
Malaysia	Maliau Seraya	Red	15	0.24	68
Ghana	Kogyae 04	NIR	14	0.23	299
Ghana	Ankasa 01	NIR	11	0.22	353
Ghana	Bobiri Y53	Green	18	0.22	124
Ghana	Ankasa 03	NIR	9	0.20	391
Malaysia	E	NIR	8	0.16	286
Malaysia	Maliau Belian	Green	17	0.12	33
Malaysia	Maliau Belian	NIR	17	0.06	143
Malaysia	Danum 01	Green	12	0.03	96
Malaysia	Danum 01	NIR	12	0.00	208
Ghana	Ankasa 01	Green	11	0.00	140
Malaysia	Danum 01	Red	12	0.00	81
Malaysia	Maliau Belian	Red	17	0.00	31

Table S7.2*:* Variance in Sentinel-2 monthly averaged reflectance from **Figure 7.5b** between plot pairs computed as a proportion of the overall variance for that GEM site (Brazil/Ghana/Malaysia). Based on the premise that reflectance should vary less between two nearby plots than across all the plots within a country. Proportion of variance results are ranked from smallest to largest and coloured by Sentinel-2 band (Green = Band 2, Red= Band 4 and Grey = Band 8).

Site	Plots	Proportion of variance	Band
Ghana	Ankasa 01+03	0.0024	NIR
Borneo	Danum 01+02	0.0088	NIR
Ghana	Ankasa 01+03	0.0094	Red
Borneo	B North+South	0.0122	Red
Ghana	Bobiri 01+02	0.0151	NIR
Ghana	Bobiri 01+02	0.0221	Red
Ghana	Bobiri 01+02	0.0229	Green
Borneo	Maliau Belian+Seraya	0.0325	NIR
Borneo	B North+South	0.0370	NIR
Borneo	B North+South	0.0465	Green
Borneo	Danum 01+02	0.0480	Green
Ghana	Ankasa 01+03	0.0927	Green
Ghana	Kogyae 02+03	0.1014	Green
Brazil	Nova Xavantina 01+02	0.1119	NIR
Borneo	Danum 01+02	0.1259	Red
Borneo	Maliau Belian+Seraya	0.1464	Green
Ghana	Kogyae 02+03	0.2950	NIR
Borneo	Maliau Belian+Seraya	0.3353	Red
Ghana	Kogyae 02+03	0.4386	Red
Brazil	Nova Xavantina 01+02	0.5860	Green
Brazil	Nova Xavantina 01+02	0.8721	Red

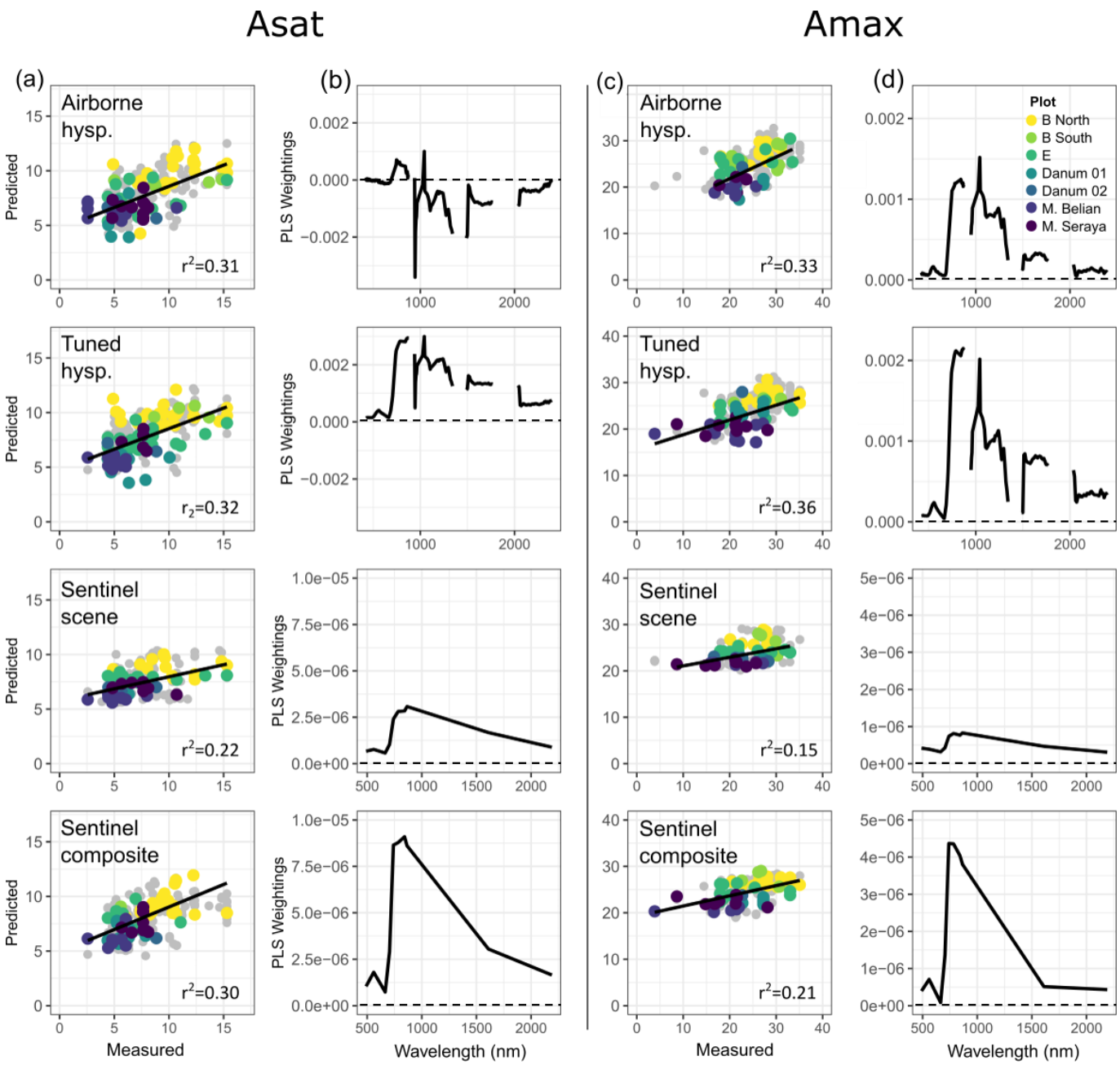


Figure S7.4*:* (a&c) Relationships between predicted versus measured A_{sat} and A_{max} values using PLSR analysis. (b&d): PLSR spectral weightings for A_{sat} and A_{max}. Deviation from 0 indicates the parts of the spectrum that contribute most strongly to prediction. Values were predicted four times using: (1) the airborne hyperspectral imagery (2) the hyperspectral imagery tuned with the Sentinel composite (**Figure 7.7**), (3) a single cloud-free Sentinel scene taken on 23rd of September, 2019 (4) a Sentinel time-series composite.

Asat	r^2	*%RMSE*	*RMSE*
Airborne hyperspectral	0.31	31	2.43
Tuned hyperspectral	0.32	31	2.44
Sentinel scene	0.22	33	2.47
Sentinel composite	0.30	33	2.53
Amax			
Airborne hyperspectral	0.33	18	4.39
Tuned hyperspectral	0.36	18	4.46
Sentinel scene	0.15	23	5.33
Sentinel composite	0.21	20	4.88

Table S7.3*:* PLSR results of predicted *versus* measured plant photosynthesis values (A_{sat} and A_{max}). Evaluated using r^2, RMSE as a percentage of the sample mean, and raw RMSE (μmol m^{-2} s^{-1}). Spectra were analysed four times using: (1) the airborne hyperspectral imagery (2) the hyperspectral imagery tuned with the Sentinel composite (**Figure 7.7**), (3) a single cloud-free Sentinel scene taken on 23rd of September, 2019 (4) a Sentinel time-series composite. All values are the average of 100 PLSR model runs.

	Spectral coverage 30%			Spectral coverage 80%		
$Vcmax_{25}$	r^2	*%RMSE*	*RMSE*	r^2	*%RMSE*	*RMSE*
Hyperspectral	0.33	32	6.56	0.27	27	6.79
Adj. hyperspectral	0.34	32	6.44	0.29	33	6.6
Sentinel scene	0.27	35	6.63	0.29	34	6.75
Sentinel composite	0.27	35	6.92	0.29	35	6.81
$Jmax_{25}$						
Hyperspectral	0.35	19	26.25	0.35	21	28.40
Adj. hyperspectral	0.34	20	27.14	0.36	20	27.00
Sentinel scene	0.18	22	29.19	0.22	23	30.69
Sentinel composite	0.21	22	29.65	0.18	22	32.34
Asat						
Hyperspectral	0.30	32	2.49	0.23	33	2.62
Adj. hyperspectral	0.30	31	2.45	0.24	33	2.60
Sentinel scene	0.19	33	2.47	0.22	34	2.52
Sentinel composite	0.27	33	2.54	0.28	35	2.65
Amax						
Hyperspectral	0.30	17	4.31	0.28	19	4.72
Adj. hyperspectral	0.30	18	4.44	0.33	18	4.49
Sentinel scene	0.14	21	5.09	0.19	21	5.09
Sentinel composite	0.22	19	4.69	0.21	22	5.27

Table S7.4: PLSR results of predicted *versus* measured plant photosynthesis traits ($Vcmax_{25}$, $Jmax_{25}$ A_{sat} and A_{max}) for tree crowns 30% covered in spectral information and 80% covered in spectral information. Evaluated using r^2, RMSE as a percentage of the sample mean, and raw RMSE (μmol m^{-2} s^{-1}). Spectra were analysed four times using: (1) the airborne hyperspectral imagery (2) the hyperspectral imagery tuned with the Sentinel composite (**Figure 7.7**), (3) a single cloud-free Sentinel scene taken on 23rd of September, 2019 (4) a Sentinel time-series composite. All values are the average of 100 PLSR model runs.

Chapter 8:

Global mapping of the photosynthetic capacity of tropical forests



Photo: Yadvinder Malhi
Wayqecha, Peru

8.1: Chapter Statement

This study uses Sentinel-2 imagery combined with environmental datasets, to create a pan-tropical map of photosynthesis capacity ($Vcmax_{25}$ and $Jmax_{25}$). Similarly to **Chapter 7**, it builds on **Chapters 5&6** by advancing a focus from mapping biochemical leaf traits to mapping process-based traits.

Following a promising regional analysis in **Chapter 7** using Sentinel-2 imagery to quantify $Vcmax_{25}$ and $Jmax_{25}$, **Chapter 8** integrates data from 35 tropical forest plots, representing all tropical forest continents, to expand predictive coverage to the whole tropical forest biome. It incorporates the finding and suggestions made in **Chapter 7**, such as using a Sentinel time-series composite in analyses, rather than single Sentinel scenes. To further strengthen predictive models, it combines Sentinel-2 imagery with environmental datasets, to reduce the influence of artefacts associated with Sentinel imagery and harness the predictive power of environmental drivers of $Vcmax_{25}$ and $Jmax_{25}$.

Chapter 8 examine patterns and magnitudes of $Vcmax_{25}$ and $Jmax_{25}$ at inter-continental, regional and local scales. It distinguishes between lowland and lowland-montane tropical forest, and draws inferences on the mechanistic relationships between photosynthetic capacity and environmental variables.

To the best of our knowledge, this is the first research that uses Sentinel-2 imagery to create a pan-tropical map of $Vcmax_{25}$ and $Jmax_{25}$.

8.2: Author Contributions

The ideas, research questions, study design, data collation, data processing, data analysis and writing presented in this chapter are the work of the first author, Eleanor Thomson. These activities greatly benefited from the contributions of Yadvinder Malhi and Chris Doughty.

GEM plot establishment and leaf photosynthesis collections are the result of over 10 years of work building the GEM network, and were carried out by a number of people, including Yadvinder Malhi, Imma Oliveras, Alexander Shenkin, Lisa Patrick Bentley, Cécile Girardin, Sabine Both, Terhi Riutta, Erika Berenguer and Chris Doughty, with extensive collaboration with in-country institutions and individuals. Tree crown maps were created by Martin Svátek.

8.3. Abstract

Plant photosynthetic capacity is determined by the maximum carboxylation rate ($Vcmax_{25}$) and the maximum electron transport rate ($Jmax_{25}$) at 25°C. Although these traits are known to vary considerably in space and time, they are typically parameterized in terrestrial biosphere models with one value for each plant functional group. Tropical forests account for one-third of terrestrial productivity but face a multitude of threats from climate change and direct human impacts. Landscape level maps of photosynthetic capacity would help determine pan-tropical drivers of photosynthetic capacity and generate robust projections of regional and global environmental change. In this study, we use a unique dataset of photosynthesis, collected from 990 trees using standardised protocols, to map $Vcmax_{25}$ and $Jmax_{25}$ across the tropics. Our field sites represent 35 tropical forest plots, cover all tropical continents, and include an elevation, precipitation, and logging gradient. Photosynthetic data was geolocated using LiDAR-mapped tree crowns, and combined with Sentinel-2 imagery and environmental datasets in Google Earth Engine to create Random Forest maps of $Vcmax_{25}$ and $Jmax_{25}$ across lowland only and lowland-montane tropical forest. Our models captured 38% and 57% of variation in $Vcmax_{25}$ and $Jmax_{25}$ across lowland only forest, and 62% and 61% of variation across lowland-montane forest. In lowland only forest, the most important predictor of $Vcmax_{25}$ and $Jmax_{25}$ was mean climatic water deficit (MCWD) and solar radiation respectively. In lowland-montane forest, the strongest predictor of photosynthetic capacity was mean annual temperature (MAT). At intercontinental scales, we find that $Vcmax_{25}$ and $Jmax_{25}$ are mainly controlled by climate. At the regional level, trait distributions are shaped by soil type. At local levels, trait patterns reflect spectral variation arising from topography and canopy structure. Our approach offers new opportunities to measure and monitor tropical forest productivity and improve the representation of photosynthesis in the next generation of Earth System Models.

Key words: photosynthetic capacity, Vcmax, Jmax, tropical forest, remote sensing, Sentinel 2, Random Forest, map.

8.4. Introduction

Humid tropical forests and savannahs are one of the most productive biomes on earth, accounting for over 60% of terrestrial net primary productivity and controlling carbon and water fluxes from regional to global scales (Foley *et al*, 2003; Beer *et al*. 2010). Yet, these ecosystems are undergoing rapid transformation as a result of environmental change and direct human impacts (Laurance *et al*. 2013; Malhi *et al*. 2014; Malhi *et al*. 2020). One of the major questions in Earth Systems Science over the last few decades has been how anthropogenic activities will affect tropical ecosystems and in turn, how altered ecosystem composition and functioning might affect global climate (Nobre *et al*. 1991; Salazar *et al*. 2007; Bonan *et al*. 2008; Mitchard *et al*. 2018).

Terrestrial biosphere models (TBMs) are the principal means for generating regional and global projections of terrestrial carbon cycling. Photosynthesis processes drive 90% of biosphere-atmosphere carbon fluxes through the conversion of solar energy to chemical energy (Joiner *et al*. 2011). Thus, representative photosynthesis data are required in order to accurately quantify the feedbacks between tropical forest productivity and environmental change. Two main leaf traits ($Vcmax_{25}$ and $Jmax_{25}$) are used to parametrize empirical models of photosynthesis, including the widely-adopted Farquhar-von Caemmerer—Berry model (Farquhar *et al*. 1980). $Vcmax_{25}$ represents light-saturated conditions at 25°C, where photosynthesis is limited by the maximum carboxylation rate of RuBisCo, while $Jmax_{25}$ represents the light-limited state at 25°C, where photosynthesis is limited by electron transport. Sensitivity analyses have shown model estimates of GPP and NPP are particularly sensitive to $Vcmax_{25}$ values (Knorr and Heimann, 2001; Zaehle *et al*. 2005; Friend *et al*. 2010), with parameter uncertainty over $Vcmax_{25}$ found to result in ~30 Pg C variation in global terrestrial GPP estimates per year (Bonan *et al.* 2011). Thus, there is considerable need to improve and constrain parametrization of $Vcmax_{25}$ and $Jmax_{25}$ in TBMs to produce robust projections of global change.

Within TBMs, $Vcmax_{25}$ and $Jmax_{25}$ data are typically sourced from the literature, or databases, and partitioned into two values for the moist tropics – one for tropical

evergreen trees and one for tropical deciduous trees (Rogers, 2014). Unsurprisingly, such generalisation fails to capture the magnitude, regional variation, and seasonality of photosynthetic capacity across tropical forest ecosystems, with obvious implications for vegetation and climate modelling. $Vcmax_{25}$ has been found to vary considerably within and between vegetation species depending on environmental and seasonal controls (Groenendijk *et al*, 2011; Osuna *et al*. 2015). Within the same plant functional group, $Vcmax_{25}$ has been found to vary by a factor of 2-3, with variation particularly pronounced in ecosystems with a large seasonal range (e.g. deciduous forests), or regions prone to drought (Dillen *et al*. 2011; Medvigy *et al*. 2013).

Progress in parametrising TBMs with accurate photosynthetic capacity values has been hindered by the difficulty in understanding and mapping $Vcmax_{25}$ at global scales (Kattge *et al*. 2009). Trait data is particularly lacking in tropical ecosystems (Jetz *et al*. 2016). Tall trees, high species diversity and the labour-intensive nature of deriving $Vcmax_{25}$ and $Jmax_{25}$ information from *A*-Ci-curves, combined with the inaccessibility of many tropical regions, means collecting $Vcmax_{25}$ and $Jmax_{25}$ data across tropical areas is a time-consuming and expensive process. In addition, differences in measurements protocols and temperature corrections often result in incompatible datasets.

Remote sensing would be an ideal method of generating spatially-continuous, consistent $Vcmax_{25}$ and $Jmax_{25}$ data. Using 'greenness' indices as a proxy for photosynthesis, MODIS offers a daily GPP and annual NPP product (Running *et al*. 2004). However, these measurements have been found to saturate over tropical forests and do not provide insight into possible variation in underlying photosynthetic mechanisms (Turner *et al*. 2006; Propastin *et al*. 2012; Cleveland *et al*. 2015). Recently, solar induced fluorescence (SIF) has been used as a more accurate proxy for GPP measurements across forest ecosystems (Guanter *et al*. 2007; Meroni *et al*. 2009), but similarly to MODIS, SIF cannot provide the photosynthetic capacity information required for TBMs.

Imaging spectroscopy has emerged as a more promising method for generating large-scale photosynthetic capacity data. At the leaf- level, numerous studies have used hyperspectral data (VNIR (400–1100) or VSWIR (400–2500 nm)) to develop predictive models of physical leaf traits, such as leaf mass per area, macro- and micronutrient

concentrations, chlorophylls, lignin, phenols and water content (Asner and Martin *et al*. 2008; Serbin *et al*. 2014; Asner *et al*. 2015; Doughty *et al*. 2017; Thomson *et al*. 2018). Using the same method, a few studies have attempted to measure $Vcmax_{25}$ and $Jmax_{25}$ directly with promising results. These analyses are based on the premise that photosynthetic processes are influenced by physical leaf characteristics, such as light-absorbing pigments and nutrient concentrations, which in turn affect the optical properties of vegetation. Highest prediction accuracies of $Vcmax_{25}$ and $Jmax_{25}$ (r^2 = 0.88-0.92) have been achieved in crop studies (Serbin *et al.* 2012; Ainsworth *et al*. 2014; Heckmann *et al*. 2017), with models found to be robust to short-term variation in $Vcmax_{25}$ and $Jmax_{25}$ across a growing season (Serbin *et al*. 2012; Ainsworth *et al*. 2014; Barnes *et al*. 2017). Across forest ecosystems, Doughty *et al*. (2011) obtained lower prediction accuracies of $Vcmax_{25}$ and $Jmax_{25}$ ($r^2 \leq 0.52$) for 11 tropical tree species, while Dechant *et al*. (2017), reported higher estimates (r^2 = 0.64-70) for 37 temperature deciduous tree species.

Scaling from the leaf to the canopy level, Serbin *et al*. (2015) used high-altitude AVIRIS (Airborne Visible-Infrared Imaging Spectroscopy) imagery to achieve very good prediction results (r^2 = 0.90) of $Vcmax_{25}$ for nine crop species. High-resolution hyperspectral imagery has been used to map physical leaf traits across tropical forests (e.g. Asner *et al*. 2016; Asner *et al*. 2017; Swinfield *et al*. 2018), suggesting that airborne hyperspectral data could be used to derive spatially-continuous $Vcmax_{25}$ and $Jmax_{25}$ values across tropical ecosystems. Currently however, hyperspectral imagery is not widely available over the tropics. While recently launched hyperspectral satellites such as NASA's HypspIRI (Lee *et al*. 2015) offer new possibilities to retrieve photosynthetic information, low spatial resolution (>60m) may create challenges in scaling up field-measured traits, especially across tropical regions (see below).

Landsat imagery has been used to map $Vcmax_{25}$ through relationships with leaf chlorophyll and nitrogen concentrations at local scales (Croft *et al*. 2016; Houborg *et al*. 2013). Similarly, vegetation indices derived from Landsat and MODIS data have been used to estimate photosynthetic capacity parameters in temperate and alpine environments (Jin *et al*. 2012; Gao *et al*. 2014). Challenges unique to tropical ecosystems however have hampered the widespread development of satellite-based photosynthetic

mapping methods across the tropics. High spatial resolution imagery is required to overcome spatial mismatches between a highly diverse tropical forest canopy (typically sampled at the level of an individual tree crown) and imagery pixel size. In addition, high revisit times are required to create cloud-free composites, especially over humid tropical forests, where clear days can be infrequent and levels of haze are high (Yanhon *et al*. 1996; Asner, 2001, Lee *et al*. 2006; Sanchez *et al.* 2020).

Sentinel-2 data from the European Space Agency's (ESA) Copernicus programme currently represents the highest resolution, openly-accessible satellite imagery, and may offer the best potential to retrieve photosynthetic information remotely. Sentinel-2 imagery has a ground sampling distance of 10m in the VNIR and 20m in the SWIR (less than the diameter of many tropical forest tree crowns), a revisit time of 5 days and 13 spectral bands. Despite capturing fewer than 10% of the spectral bands measured by a hyperspectral sensor, the majority of Sentinel-2 spectral channels are clustered in the visible, red-edge and near-infrared parts of the spectrum - regions that contain many absorption features linked to specific leaf traits and are thus disproportionally important in measuring these traits (Curran, 1989; Kokaly *et al*. 2009; Ustin *et al*. 2009). Accordingly, Sentinel-2 imagery has been used to map pan-tropical physical and chemical leaf traits (Aguirre-Gutiérrez *et al*. 2020), as well as biochemical leaf traits in tropical, temperate and Arctic environments (Gara *et al*. 2019; Ma *et al*. 2019; Thomson *et al*. 2021). To the best of our knowledge however, this is the first research that uses Sentinel-2 imagery to create pan-tropical maps of $Vcmax_{25}$ and $Jmax_{25}$.

In this study, we use a large, unique and consistent dataset of photosynthetic traits and forest inventory data, collected using standardised protocols across 35 tropical forest plots in the Global Ecosystems Monitoring (GEM) network (Malhi *et al.* 2021), to create predictive models of $Vcmax_{25}$ and $Jmax_{25}$. The 35 tropical forest plots are from six main study sites covering all four tropical continents and span temperature, precipitation and logging gradients. A complementary dataset of geolocated tree stems and LiDAR-mapped tree crowns are used to georeference individual tree crowns. Trait data was combined with Sentinel-2 imagery and environmental datasets in Google Earth Engine to create Random Forest maps of $Vcmax_{25}$ and $Jmax_{25}$ across lowland and lowland-montane tropical forest.

The study has three aims:

1. To investigate how photosynthetic capacity ($Vcmax_{25}$ and $Jmax_{25}$) varies across six tropical forest sites and across environmental gradients at each site.
2. To determine which spectral and environmental variables are most important in predicting $Vcmax_{25}$ and $Jmax_{25}$ information pan-tropically.
3. To examine how the magnitude and spatial patterns of $Vcmax_{25}$ and $Jmax_{25}$ vary within and between lowland and montane tropical forest ecosystems.

8.5. Materials and Methods

Study Sites

Vegetation data were collected from 35 permanent tropical forest plots located at six sites across the tropics: Australia, Brazil, Gabon, Ghana, Malaysia and Peru. All plots are part of the Global Ecosystems Monitoring (GEM) network (Malhi *et al.* 2021). The Australian and Peru plots are along temperature gradients, extending from humid tropical forest in the lowlands to tropical montane forest at 1140 m.a.s.l in Australia and 3537 m.a.s.l in Peru. The plots in Ghana are along a precipitation gradient stretching from wet evergreen tropical forest, through semi-deciduous tropical forest, to dry forest-savannah transition vegetation. The Malaysian plots are located in Sabah, Borneo and, along with the Brazilian plots, span a forest disturbance gradient from old-growth primary forest to logged forest fragments. Hereafter, the Brazilian plots are referred to Brazil-ST to distinguish them from the Nova Xavantina GEM site in Brazil. The Gabon plots represent a defaunation contrast, from impoverished coastal forest to the mega-diverse Lopé National Park, home to a large population of forest elephants and other megafauna. With the exception of the Brazil-ST plots (which are 0.25 ha), all plots are 1 ha. Within each plot, multiple forest censuses have been undertaken, with all trees >10cm diameter at breast height (DBH) tagged and identified at species level. See **Table 8.1** and *8. 10: Supplementary material* for more details.

Plot ID	Region	Country	Forest type	Latitude	Longitude	Size (ha)	Elevation (m.a.s.l)	Data of photosynthesis collection	Spatial fold group
AEP02	Australian Tropics	Australia	Submontane forest	145.586	-17.146	0.5	710	July-September, 2015	1
AEP03			Tropical montane cloud forest	145.592	-17.088	0.5	1140		2
AEP33			Submontane forest	145.572	-17.286	0.5	730		1
AEP41			Lowland moist forest	145.441	-16.136	0.5	15		3
DRO01				145.447	-16.103	1	58		3
ROB06			Submontane forest	145.630	-17.121	1	720		1
ST0-03	Amazonia	Brazil	Lowland moist forest	-54.964	-3.297	0.25	133	August-September, 2015	4
STQ-08			Logged	-54.675	-2.515	0.25	135		5
STQ-11				-54.671	-2.574	0.25	114		5
LPG-01	Central Africa	Gabon	Lowland moist forest	11.574	-0.174	1	399	February-March, 2017	6
LPG02				11.615	-0.216	1	288		6
MNG04				9.324	0.577	1	39		7
ANK-01	West Africa	Ghana	Lowland moist forest	-2.695	5.268	1	114	December 2015-March 2016	8
ANK-03				-2.692	5.271	1	86		8
BOB-01				-1.319	6.705	1	268		9
BOB-02				-1.338	6.691	1	276		9
KOG-01			Forest-savannah transition	-1.150	7.262	1	230	October, 2015	10
KOG-03				-1.156	7.307	1	196		10
KOG-04				-1.180	7.303	1	227		10
DAN-04	Borneo	Malaysia	Lowland moist forest	117.796	4.952	1	238	July-December, 2015	11
DAN-05				117.794	4.953	1	230		11
MLA-01				116.970	4.747	1	524		12
MLA-02				116.950	4.754	1	493		12
SAF-01			Logged	117.618	4.732	1	388		13
SAF-02				117.618	4.739	1	580		13
SAF-03				117.588	4.692	1	170		13
ACJ-01	Tropical Amazon-Andes	Peru	Tropical montane cloud forest	-71.632	-13.147	1	3537	April-November, 2013	14
ESP-01				-71.595	-13.176	1	2863		15
PAN-02			Submontane forest	-71.263	-12.650	1	595		16
PAN-03				-71.274	-12.638	1	848		16
SPD-01			Tropical montane cloud forest	-71.542	-13.048	1	1776		17
SPD-02				-71.537	-13.049	1	1527		17
TAM-05			Lowland moist forest	-69.271	-12.830	1	223		18
TAM-06				-69.296	-12.839	1	215		18
WAY-01			Tropical montane cloud forest	-71.587	-13.191	1	3045		19

Table 8.1: Table of all GEM plots used in the analysis and associated characteristics. Spatial fold groups created for spatial cross-validation (see *8.5. Random Forest Analysis*). Plots located < 80km apart, and < ±350m in elevation are included in the same group.

Leaf photosynthesis collection

A_{sat}, A_{max} and dark respiration (R_{dark}) measurements were carried out using standard GEM protocols (Oliveras *et al*. 2020). Within each plot, the most abundant species were chosen that represented 60-80% of the basal area of each plot. In low diversity sites, such as savannah or montane forest, >80% of the basal area of a plot could be sampled, whereas in the higher diversity plots, only 60% basal area sampling was achieved. From the most abundant species, one to five individuals were sampled, with three to five fully expanded, undamaged sun leaves harvested from an upper-canopy branch in direct sunlight. For each leaf, A_{sat}, A_{max} and R_{dark} measurements were carried out in-situ using a LiCOR machine. 30 readings were averaged per measurement. Measurements were checked for consistency and those displaying anomalous light or CO_2 recordings removed. See Oliveras *et al*. (2020), Gvozdevaite *et al*. (2018) and Both *et al*. (2019) for more details on individual trait campaigns. Across all plots, this campaign resulted in a total sample of 990 individual trees.

Due to the time-consuming nature of their collection, *A*-Ci-curves normally used to estimate $Vcmax_{25}$ and $Jmax_{25}$ were not taken at all sites, or only collected as a subset of data. In order to generate a consistent dataset of $Vcmax_{25}$ and $Jmax_{25}$, the one-point method from the plantecophys R package (Duursma, 2015) was applied to A_{sat}, A_{max} and R_{dark} data at all sites. This method generates $Vcmax_{25}$ and $Jmax_{25}$ estimates by inverting the Farquhar photosynthesis equation, and is based on a study by de Kauwe *et al*. (2016) who used A_{sat} and R_{day} measurements to estimate $Vcmax_{25}$ with an r^2 of 0.95 and RMSE of 17.1 μmol m^{-2} s^{-1}. We tested the one-point method at GEM sites where *A*-Ci-curves were collected: Australia, Brazil and Gabon. Good agreement was found between values derived from A-Ci-curves and those estimated using the one-point method, with r^2 falling between 0.57-79 for $Vcmax_{25}$ and 0.63-92 for $Jmax_{25}$ (**Figure S8.1**).

While the trait sampling was designed to represent the most dominant species in the plot, some individual trees lacked field-sampled trait values. In order to include these trees in our dataset, representative trait values were attributed to some individuals using the following method 1) individuals from which photosynthesis measurements had been taken kept their original trait values. 2) For individuals of a species from which sampled

individuals were present in the plot, a distance-weighted mean was attributed, that reflected the trait values of the nearer sampled individual(s) more strongly than the trait values of the further sampled individual(s). A distance-weighted mean was used to account for the influence of microclimates, micro-variations in soil, and the likelihood that individuals in close proximity share genetic information. 3) If individuals of a species were not sampled in the plot, but were sampled in a nearby plot, a species mean from outside the plot was attributed. A nearby plot was regarded as one <5km away, located in the same type of forest and at the same elevation (see the spatial folds column **Table 8.1**) For individuals belonging to a species that was not measured in the vicinity, steps 2 and 3 were repeated for genera and family, with the exception that a plot-level mean was attributed in step 2, rather than distance-weighted mean. No trait values were assigned to remaining unsampled trees. Attributing taxonomic means to individuals is supported by studies that show that variation in photosynthetic traits is strongly taxonomically driven at a local scale, with species, genera and family influencing trait distributions more strongly than environmental variation (Asner *et al.* 2014; Oliveras *et al.* 2020). See **Figure S8.2** for only field-sampled trait values.

Sentinel-2 imagery

Over each site, a cloud-free composite was created in Google Earth Engine using 2019-20 Sentinel-2 Bottom-Of-Atmosphere reflectance imagery. For each site, a composite was created by taking all the available Level-2A images over the site during the photosynthesis sampling months (and one month either side if the sampling period was <3months), filtering each image to leave only clear vegetation pixels (as detected by ESA's scene classification layer (SCL)), and taking the median value of all the remaining reflectance values for each pixel in the time-series. This compositing method has been shown to provide the strongest signal-to-nose ratio for Sentinel imagery and reduce artefacts associated with remaining cloud cover, cloud shadow and haze (**Chapter 7**). SWIR bands at 20m were resampled to 10m using bilinear interpolation (**Figure 8.1(b)**).

While in some cases, top-of-atmosphere Level-1C Sentinel imagery was available in years closer to the original photosynthesis collection, these images required atmospheric correction using the Sen2Cor processor. Downloading and processing the number of scenes required for a cloud-free composite over several sites in the tropics with the Sen2Cor processor requires considerable computing storage and resources. Thus, ready-computed 2019 Level-2A products in Google Earth Engine were used instead. It is expected that the composition and structure of undisturbed tropical forest plots would not have changed substantially in 4-5 years. During this time, the forest censuses show little change in species distribution or basal area (GEM, unpublished data). To further evaluate the use of 2019 composite imagery, all analyses were re-computed with single Sentinel scenes taken as close as possible to the original trait collection. See *8.10: S*upplementary material for results.

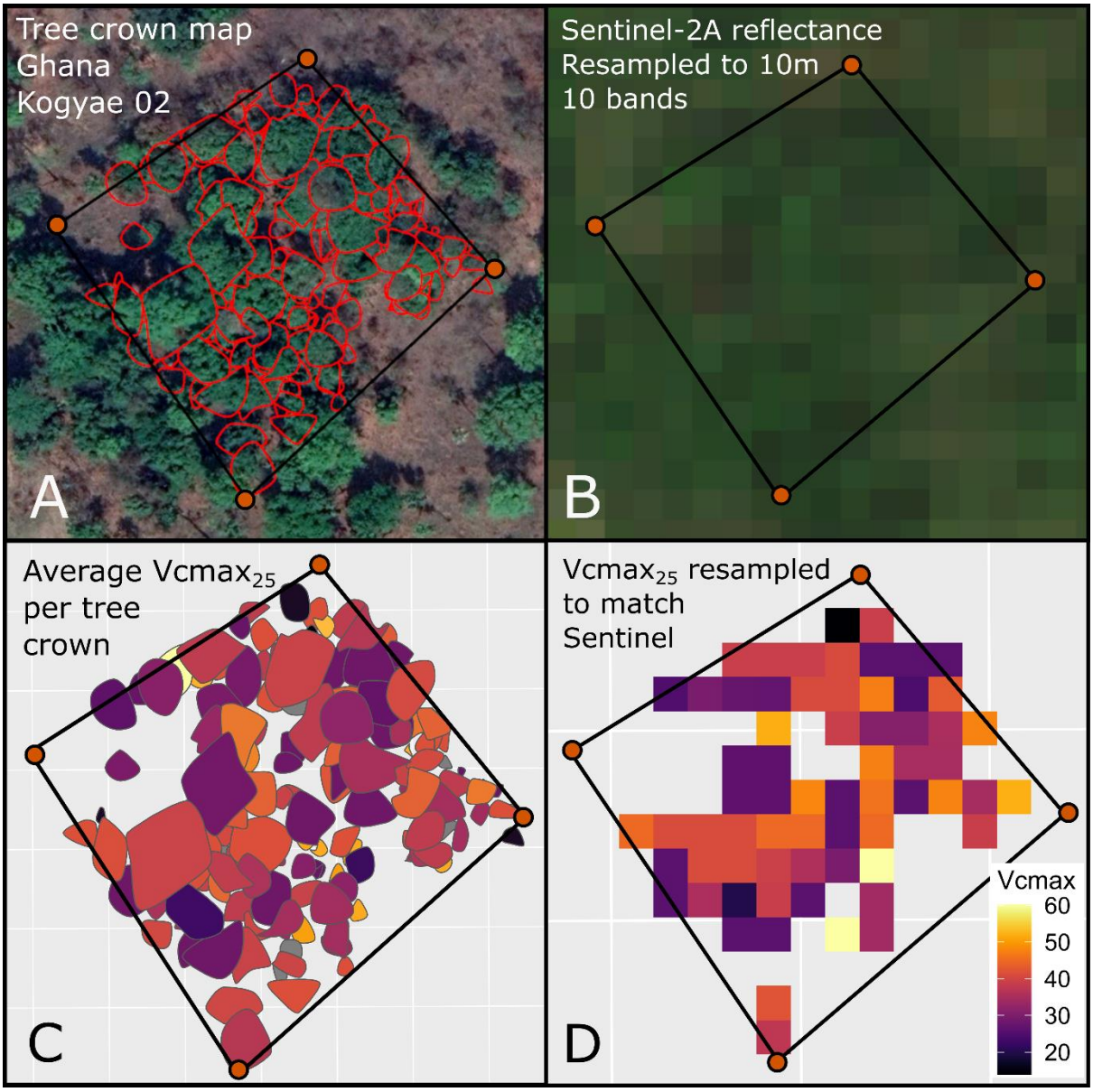


Figure 8.1: An example of the approach taken to match field measurements to Sentinel-2 imagery (A) Very high-resolution imagery of Kogyae 02 (taken from Google Earth: accessed 24/08/2021) overlain with plot corners and the laser tree crown map (B) Sentinel-2 imagery overlain with Kogyae 02 plot corners. (c) Tree crowns coloured with corresponding $Vcmax_{25}$ values. (D) $Vcmax_{25}$ values resampled to match resolution and extent of Sentinel-2 imagery. Only pixels >60% covered by tree crowns with trait information were included.

Tree crowns maps

Tree crown maps were used to geolocate individual tree stems and tree crown outlines over Sentinel imagery. Tree crown maps were generated using two methods. For the Ghana, Malaysia and Peru plots (66% of plots), stem and crowns positions were recorded to decimetre precision using the ground-based laser method (Field-Map technology, IFER, Czech Republic, https://www.field map.cz/) described in Hédl *et al*. (2009). Between 5 and 30 points were recorded for each crown and fitted with smoothed spatial polygons. The crown polygons were then georeferenced to the plot origins using a Geneq SXBlue II GPS unit, which uses satellite-based augmentation to perform differential correction and provides subcanopy positional accuracy of less than 1m.

For the Australia, Brazil-ST and Gabon plots, local crown maps were not available. Instead, regional level allometric models developed by Shenkin *et al*. (2020) were used to estimate crown diameter. The equations are based on tree DBH and provide specific scaling factors for lowland, montane and savannah tropical forests Individual tree stems were geolocated using *xy* coordinates and GPS-located plot corners. Crown shape was assumed to be circular.

In both approaches, height measurements were used to filter overlapping tree crowns, leaving only the tallest sun-crowns that determine plot reflectance (**Figure 8.1(a)**). These crowns were matched with their mean field-measured, or attributed, photosynthesis value (**Figure 8.1(c)**).

To overcome the spatial mismatch between differing crown sizes and a 10m pixel size, the tree crown maps with their associated trait values was resampled using bilinear interpolation to match the resolution and spatial extent of the Sentinel-2 imagery (**Figure 8.1(d)**). This resampling created a community-weighted trait value for each 10m Sentinel pixel. Only pixels that were at least 60% covered by trait information were included in the analysis. Overall, this method resulted in a sample size of 360 pixels for Australia, 44 for Brazil-ST, 187 for Gabon, 468 for Ghana, 194 for Malaysia and 572 for Peru.

Model predictors

Spectral information

10 Sentinel bands were used in the analysis ((Band 2 (492 nm), 3 (560 nm), 4 (665 nm), 5 (704 nm), 6 (741 nm), 7 (783 nm), 8 (833 nm), 8A (865 nm), 11 (1614 nm) and 12 (2202 nm)). Sentinel bands 1, 9 and 10 were omitted, as these bands are used for aerosol and cloud detection and are not designed for ecological mapping. In addition to the raw bands, five common vegetation indices were computed: GCI, NDRE, NDVI, MCARI and MSAVI (see *Table 2* for more information). These indices are designed to accentuate particular properties of vegetation, such as chlorophyll and nutrient contents, and have been found to show relationships with photosynthetic capacity (Jin *et al*. 2012; Muraoka *et al*. 2013; Ide *et al*. 2010). Therefore, we hypothesised that including vegetation indices may offer more predictive capability than single spectral bands alone.

Texture metrics

Canopy structure plays an important role in determining forest functional diversity (Kamoske, 2021) and has been found to influence trait distributions across tropical forests (Aguirre-Gutiérrez *et al*. 2020). In order to characterize canopy structure, we used texture metrics computed using the Grey Level Co-Occurrence Matrix (GLCM) developed by Haralick *et al*. (1973). This method uses a specified kernel size (in this case a 3x3 pixel window) to calculate how often pairs of pixels with specific values, and in a specified spatial relationship, occur within the window. Statistical metrics are extracted as measures of texture. We calculated two texture metrics; correlation (a measure of patterning within a window) and entropy (a measure of the uniformity of pixel values within a window). These metrics were measured on three Sentinel bands: Band 2 (Green), Band 3 (Red) and Band 4 (NIR). These bands were chosen as they are particularly sensitive to vegetation properties and are measured at 10m resolution.

Climate and soil information

In addition to spectral information, climate and soil data were extracted for each pixel, to capture mechanistic relationships between environmental variables and photosynthesis capacity, and to provide insight into the abiotic controls determining $Vcmax_{25}$ and $Jmax_{25}$ at a pan-tropical scale.

Gridded climate data were extracted from TerraClimate (Abatzoglou *et al.* 2018) and averaged over a 30-year period (1990-2021). Three climate variables were included in the analysis; maximum annual temperature (MAT), annual maximum climatic water deficit (MCWD) and solar radiation (SRAD). Soil data were extracted from ISRIC SoilGrids (de Sousa *et al*. 2020) at 250m resolution. Values were averaged for the top 30cm of soil. Six soil parameters were used in the analysis; %sand, %clay, soil pH, cation exchange capacity, nitrogen content and soil organic carbon content. Lastly, slope data (Farr *et al*. 2007) were used to characterise plot topography, which has been shown to affect tropical forest leaf chemistry and functioning (Swinfield *et al*. 2019; Schmitt *et al*. 2020).

Although the pixel resolution of the environmental data was lower than that of the spectral and trait data (**Table 8.2**), we do not believe that the spatial mismatch introduced issues of spatial autocorrelation into our dataset (see discussion). In order to fully exclude spatial autocorrelation effects however, all analyses were validated using a spatial fold method, which ensures that no pixels from a plot (or nearby plot) used in model calibration entered the validation dataset (see Random Forest analysis below).

Type	Predictor	Spatial Resolution (m)	Description	References
Sentinel-2 surface reflectance	B2 (490), B3 (560), B4 (665), B8 (842)	10	Sentinel-2 bands with spatial resolution of 10m	https://sentinel.esa.int/web/sentinel/missions/sentinel-2
	B5 (705), B6 (740), B7 (783), B8A (865), B11 (1610), B12 (2190)	20	Sentinel-2 bands with spatial resolution of 20m	
Vegetation Index	GCI	10	Green Chlorophyll Index	Gitelson *et al*. (2003)
	NDRE	10	Normalized Difference Red Edge	Gitelson and Merzylak (1996)
	NVDI	10	Normalized Difference Vegetation Index	Kriegler *et al*. (1969)
	MCARI	10	Modified Chlorophyll Absorption in Reflectance Index	Daughtry *et al*. (2000)
	MSAVI2	10	Modified Soil Adjusted Vegetation Index 2	Qi *et al*. (1994)
Texture	Entropy	10	GLCM entropy calculated on Green (B3), Red (B4) and NIR (B8) bands	Haralick *et al*. (1973)
	Correlation	10	GLCM correlation calculated on Green (B3), Red (B4) and NIR (B8) bands	
Climate	MAT	4000	Mean annual temperature	Abatzoglou *et al.* (2018)
	MCWD	4000	Mean annual climatic water deficit	
	SRAD	4000	Solar Radiation	
Soil	Sand (%)	250	Amount of sand (g/kg)	de Sousa *et al*. (2020)
	Clay (%)	250	Amount of clay (g/kg)	
	Cec	250	Cation exchange capacity at pH7	
	SoilpH	250	Soil pH in H_2O	
	Nitrogen	250	Soil nitrogen	
	Carbon	250	Soil organic carbon	
Terrain	Slope	90	Terrain slope	Farr *et al*. (2007)

Table 8.2: Table of spectral, climatic and soil variables used as predictors in the Random Forest models. Sentinel bands are extracted from a cloud-free composite created in Google Earth Engine. Texture metrics are computed using the Grey Level Co-Occurrence Matrix (GLCM) method. Climate data is averaged from 1990-2021. Soil data are reported for the top 30cm of soil.

Random Forest analysis

Predictive models of $Vcmax_{25}$ and $Jmax_{25}$ were built using Random Forest regression (Breiman, 2001). Pixel-level trait values formed the dependent (target) dataset, with the spectral, climate and soil data forming the independent dataset (or predictors). Random Forest is a non-parametric machine learning classifier that grows an ensemble of decision trees. Random Forest is based on the premise that a group of classifiers perform more accurately than a single classifier (Kotsiantis and Pintelas, 2004). Each tree is independently constructed using a different bootstrap sample of the dataset ('bagging'). The data not used to grow the tree are reserved in an 'out-of-bag' (oob) subset to evaluate the performance of the tree. For each pixel, all the decision trees which did not include that pixel in their training data (around 30%) predict a trait value for the pixel, and the pixel is assigned the average of all the estimated values.

The performance of the model was evaluated using three metrics: the coefficient of determination (r^2), RMSE, and %RMSE to facilitate a comparison between traits. Each reported metric represented an average of 100 model runs. All model analyses were carried out in rstudio using the randomForest package. Parameters were tuned using the following parameters: *ntree*=500 and *mtry* =10.

In order to evaluate which predictors were most important in classifying traits, predictor importance was calculated using the 'important' function in the randomForest package. This function permutes each predictor in turn and records the loss in accuracy (mean squared error), averaged over all decision trees and normalized by the standard deviation of the difference. Importance for each predictor was recorded as a percentage of the total.

From the plot-level photosynthetic capacity data, it was clear that different controls were operating in the montane plots compared to the lowland plots. Therefore, both a lowland and a lowland-montane tropical forest model were created for $Vcmax_{25}$ and $Jmax_{25}$. All trait data were used to parameterise the lowland-montane model, whereas the lowland only model excluded data from plots >1000m (AEP-03 in Australia and WAY-01, ESP-01 and ACJ-01 in Peru).

Lowland and lowland-montane models were validated using two methods. Firstly, a standard cross-validation was used, where all pixels were treated independently and 70% were randomly sampled to create a training dataset, with 30% reserved as the validation dataset. In addition, considering the ongoing discussion around the influence of spatial autocorrelation in geospatial validation, a spatial fold method of validation was employed (Ploton *et al*. 2020). Spatial fold validation splits the pixels into spatial fold groups, which ensures that no pixels from the same plot (or plots located nearby) can be included in both the training and the validation dataset. In practice, the spatial fold method created a buffer around training pixels and excluded all pixels from the validation dataset that were located closer than 80km, or ±350m elevation to a training pixel. See **Table 8.3** for the spatial fold groups.

To explore model accuracy at a local level, lowland models were created for each site (Australia, Brazil-ST, Gabon, Ghana, Malaysia and Peru) using only the data from that site. Montane plots were included to create site-level lowland-montane models for Australia and Peru.

To quantify the accuracy of using just spectral data to predict photosynthetic capacity, models were re-run only using the predictors computed with the spectral information (e.g. raw Sentinel bands, vegetation indices and texture metrics) and excluding environmental data (see *8.10. Supplementary material* for results).

Creating pan-tropical map of $Vcmax_{25}$ and $Jmax_{25}$

To create pan tropical maps of $Vcmax_{25}$ and $Jmax_{25}$, a random forest model was created in Google Earth Engine using all applicable training pixels (e.g. all pixels for the lowland-montane model and all pixels <1000m for the lowland model). All applicable training pixels were used, on the assumption that the more information a model is trained upon, the stronger a model will perform. No pixels were kept aside for validation purposes. Each 10m pixel across the tropics was classified using the random forest model and spatially-continuous layers of spectral, textual, climate, soil and topography information described above.

To constrain the model to tropical forest pixels, a map of the tropical forest biome was downloaded from the Global Forest Monitoring project (Hansen *et al*. 2010). Within this area, only pixels containing >50% tree cover (Hansen *et al*. 2013) were classified. For the lowland tropical forest model, only pixels >1000m were classified.

A confidence map for each trait was created by taking the predictions generated by each 'tree' in the Random forest model (n=500) and calculating the standard deviation of the trait estimates for each 10m pixel.

All analyses were also carried out with A_{sat} and A_{max} data (see *8.10. Supplementary material* for results).

8.6. Results

Field-measured photosynthetic capacity varied significantly within and between tropical forest sites (**Figure 8.2**). Lowest $Vcmax_{25}$ mean values were found in Malaysia, followed by Brazil-ST (17 ± 6 and 21 ± 9 26 μmol/m^{-2} s^{-1} respectively). Highest $Vcmax_{25}$ mean values were found across Ghana and most of the Australia plots (36 ± 11 and 32 ± 10 μmol/m^{-2} s^{-1}). $Vcmax_{25}$ was significantly higher in Ghana than Gabon (29 ± 12 μmol/m^{-2} s^{-1}), despite both sites displaying similar species composition.

At the local level, the Peru elevation gradient displayed the largest $Vcmax_{25}$ range (25-100 μmol/m^{-2} s^{-1}). $Vcmax_{25}$ was strongly correlated with elevation. The lowest $Vcmax_{25}$ values were recorded in the lowland plots (<850m: 25 ± 10 μmol/m^{-2} s^{-1}), intermediate values in the mid-elevation pots (1500-1800: 36 ± 12 μmol/m^{-2} s^{-1}), high values in the high elevation plots (2800-3100m: 49 ± 15 μmol/m^{-2} s^{-1}) and extremely high values at the tree line in ACJ-01 (>3500m: 100 ± 12 μmol/m^{-2} s^{-1}). There was less of a clear $Vcmax_{25}$-elevation relationship across the Australia montane gradient, although average $Vcmax_{25}$ was lowest in the lowest elevation plot (AEP41) and highest in the highest elevation plot (AEP03). Across the Ghanaian precipitation gradient, all plots displayed relatively high $Vcmax_{25}$. Although they displayed the lowest site-level values overall, the forest disturbance plots in Malaysia recorded significantly higher $Vcmax_{25}$ than the Malaysia old-growth forests (21 ± 8 compared to 16 ± 4 μmol/m^{-2} s^{-1}). Similarly, in Brazil-ST, the secondary forest plots recorded significantly higher $Vcmax_{25}$ than the undisturbed plot.

$Jmax_{25}$ averages were similar across Australia, Brazil-ST and Gabon (68-71 μmol/m^{-2} s^{-1}), with Ghana again recording higher $Jmax_{25}$ than Gabon (**Figure 8.2**). Across the Peru elevation gradient, $Jmax_{25}$ covaried strongly with $Vcmax_{25}$, with ACJ-01 displaying the highest $Jmax_{25}$ plot average (167 ± 26 μmol/m^{-2} s^{-1}). The greatest difference in $Vcmax_{25}$/$Jmax_{25}$ ratios was shown by the Malaysian plots. In contrast to their low $Vcmax_{25}$, the Malaysian plots displayed the highest $Jmax_{25}$ averages of all the sites. Again, the logged forest plots displayed significantly higher $Jmax_{25}$ values than the old-growth plots across both Malaysia and Brazil-ST. $Jmax_{25}$ did not differ significantly across the Gabon defaunation gradient.

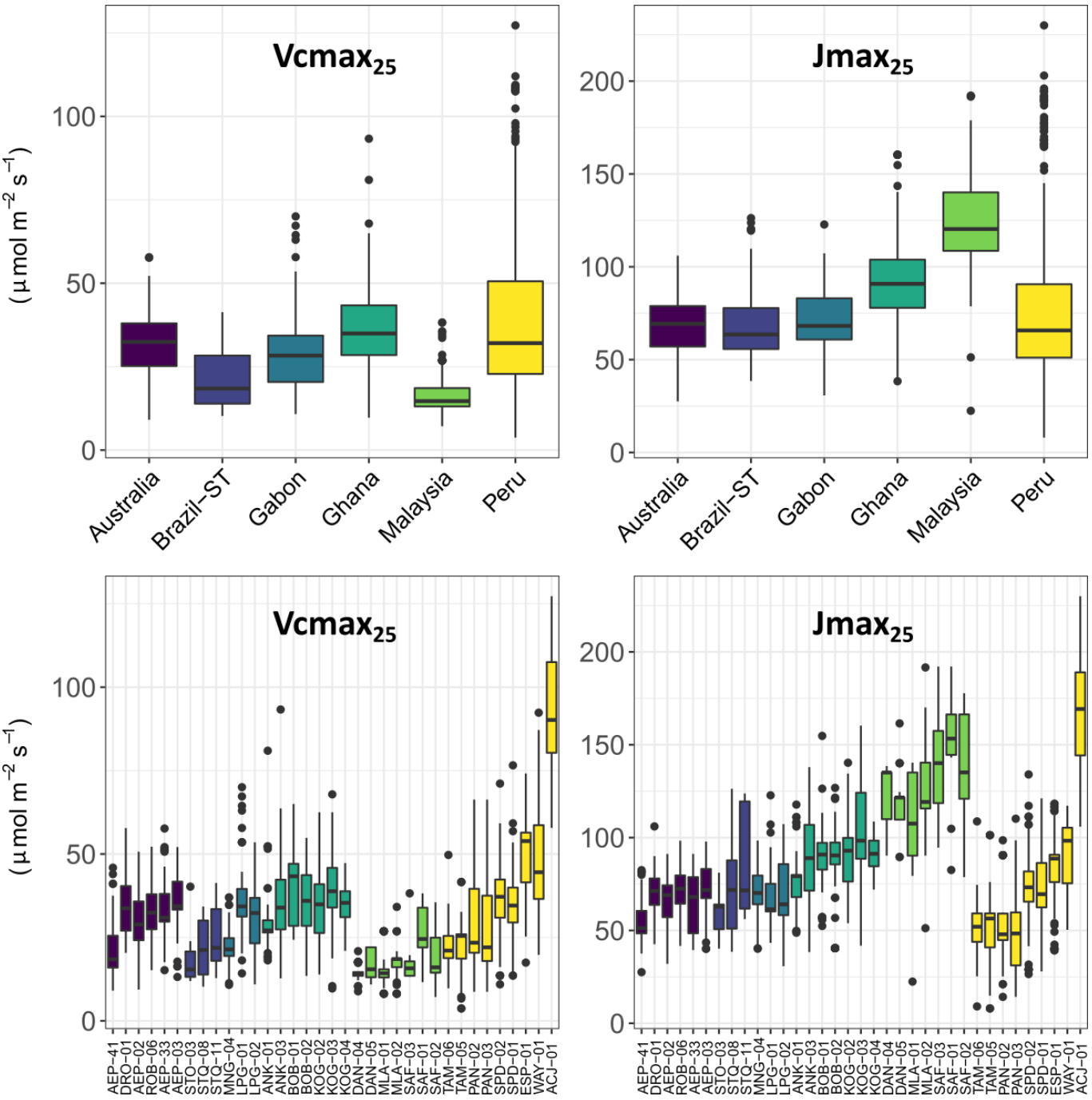


Figure 8.2: Trait values represent 10x10m pixel-level community-weighted mean of the most visible sun foliage. (*Top*) Comparison of $Vcmax_{25}$ and $Jmax_{25}$ values across 6 pan-tropical sites. (*Bottom*) Comparison of $Vcmax_{25}$ and $Jmax_{25}$ values across 35 tropical forest plots, coloured by site. More information on plot characteristics can be found in **Table 8.1**. $Vcmax_{25}$ and $Jmax_{25}$ values were derived using the one-point method (de Kauwe *et al.* 2016).

	Lowland Forest					
	$Vcmax_{25}$			$Jmax_{25}$		
	R-Squared	*RMSE*	*%RMSE*	*R-Squared*	*RMSE*	*%RMSE*
All plots	0.38	9.20	32	0.57	18.90	24
Spatial folds	0.16	11.71	40	0.30	27.52	34
Australia	0.17	8.59	28	0.15	13.02	20
Brazil-ST	0.24	8.17	38	0.29	22.66	31
Gabon	0.32	9.85	34	0.12	15.51	22
Ghana	0.09	10.12	28	0.15	20.54	23
Malaysia	0.27	5.28	32	0.26	23.69	19
Peru	0.08	9.71	40	0.06	18.78	36
	Lowland-Montane Forest					
All plots	0.62	10.28	31	0.61	19.21	24
Spatial folds	0.46	15.39	46	0.16	32.20	40
Australia	0.17	8.66	28	0.15	13.19	20
Peru	0.73	11.90	31	0.72	20.28	27

Table 8.3: Statistical results of predictive models for $Vcmax_{25}$ and $Jmax_{25}$ created using random forest regression. All statics represent the average of 100 model runs. Only data from lowlands plots (<1000m) was used to train lowland forest models. All plot data were used to train lowland-montane forest models. Pan-tropical models were validated twice. (1) using a random 70:30 training/validation split where all pixels were treated as independent samples. (2) splitting the plots into spatial folds to ensure that no pixels from the same, or nearby, plots used in the training dataset entered the validation dataset. See *Table 1* for spatial fold groupings. Models were also run at the site level using the site-specific data.

Across lowland tropical forest, Random Forest models were more accurate at predicting $Jmax_{25}$ than $Vcmax_{25}$ (r^2=0.57 compared to 0.38, **Table 8.3**). This result appears to be driven by larger site-level differences in $Jmax_{25}$ (in particular, the higher $Jmax_{25}$ values in Malaysia), creating a strong site-level $Jmax_{25}$ gradient (**Figure 8.3**). $Vcmax_{25}$ displayed less inter-site variation, with a corresponding decrease in model performance.

At the site level, local lowland models were weaker than the pan-tropical model (r^2=0.06-0.32), reflecting less intra-site than inter-site variation (**Table 8.3**). For $Vcmax_{25}$ and $Jmax_{25}$, the strongest local models were found along the forest disturbance gradients of Brazil-ST and Malaysia (r^2=0.24-27 and r^2=0.26-29 respectively). A relatively strong $Vcmax_{25}$ model was found in Gabon (r^2=0.32).

Across lowland-montane forest, pan-tropical models displayed higher prediction accuracies than lowland models alone (**Table 8.3**). Predictive performance of $Vcmax_{25}$ and $Jmax_{25}$ increased from r^2= 0. 38 and 0.57 to r^2 = 0.62 and 0.61 respectively. %RMSE did not change. At the site level, the local Peruvian lowland-montane model displayed the highest prediction accuracy of all models (r^2=0.73 and 0.72 for $Vcmax_{25}$ and $Jmax_{25}$ respectively), reflecting the strong variation in environmental conditions along the gradient.

Validating the models using two different methods ((1) a standard 70:30 cross-validation and (2) a spatial fold cross-validation) produced significantly different results (**Table 8.3**). Validation statistics were significantly lower using the spatial folds method than the standard cross-validation. Predictive performance decreased from r^2=0.38 to 0.16, and r^2= 0.62 to 0.30 for the lowland $Vcmax_{25}$ and $Jmax_{25}$ spatial folds models respectively. Spatial folds validation still produced a relatively strong lowland-montane model for $Vcmax_{25}$ (r^2=0.46), but a much weaker one for lowland-montane $Jmax_{25}$ (r^2=0.16).

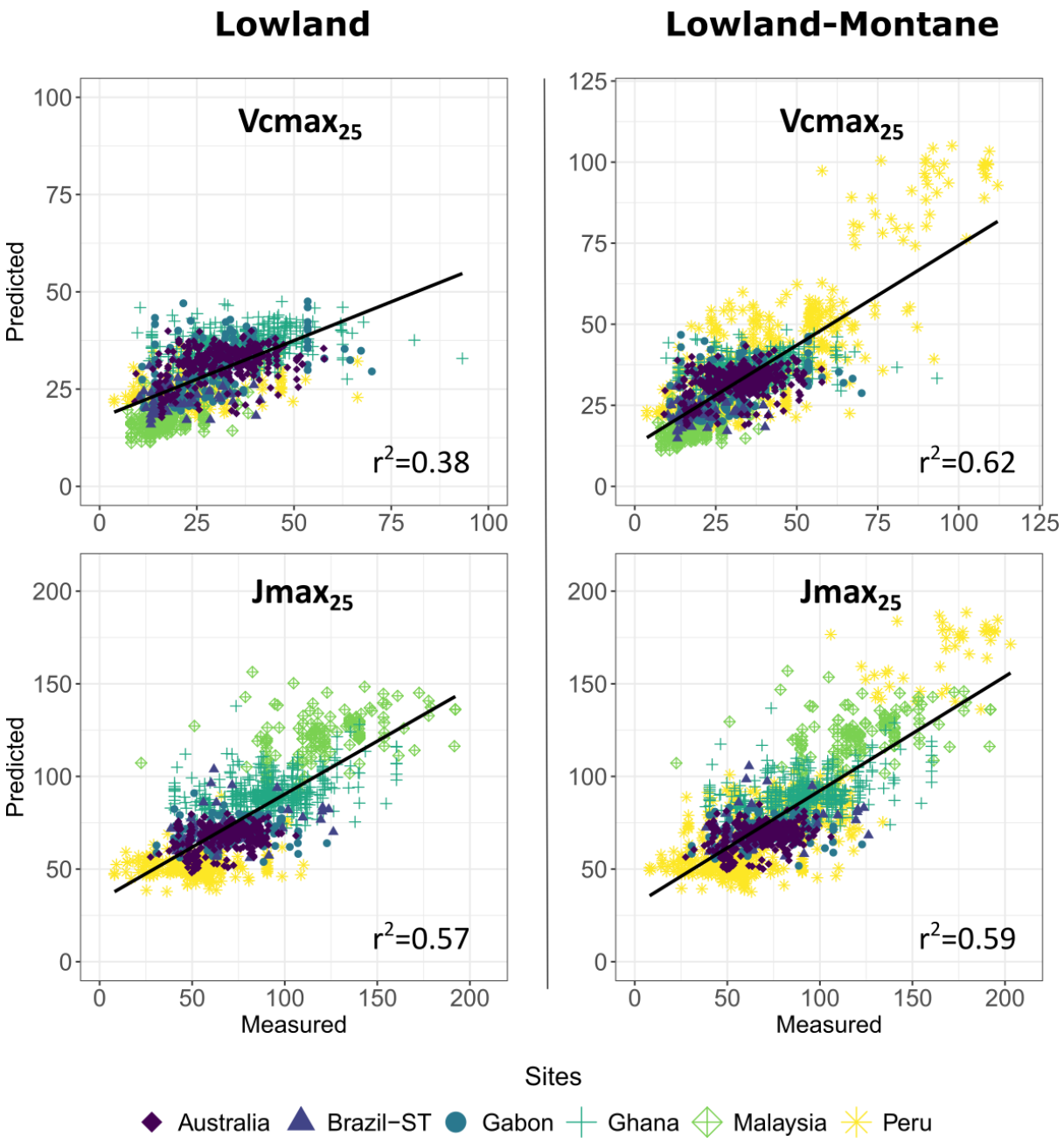


Figure 8.3: Pixel-level $Vcmax_{25}$ and $Jmax_{25}$ values plotted against model predictions for lowland and lowland-montane forest. Points represent random forest regression results for the 30% 'out of bag' validation dataset. Points are coloured according to site. The black line shows the 1:1 relationship between observed and predicted values. Model prediction accuracy is shown in the bottom left. Full model results can be found in **Table 8.3**.

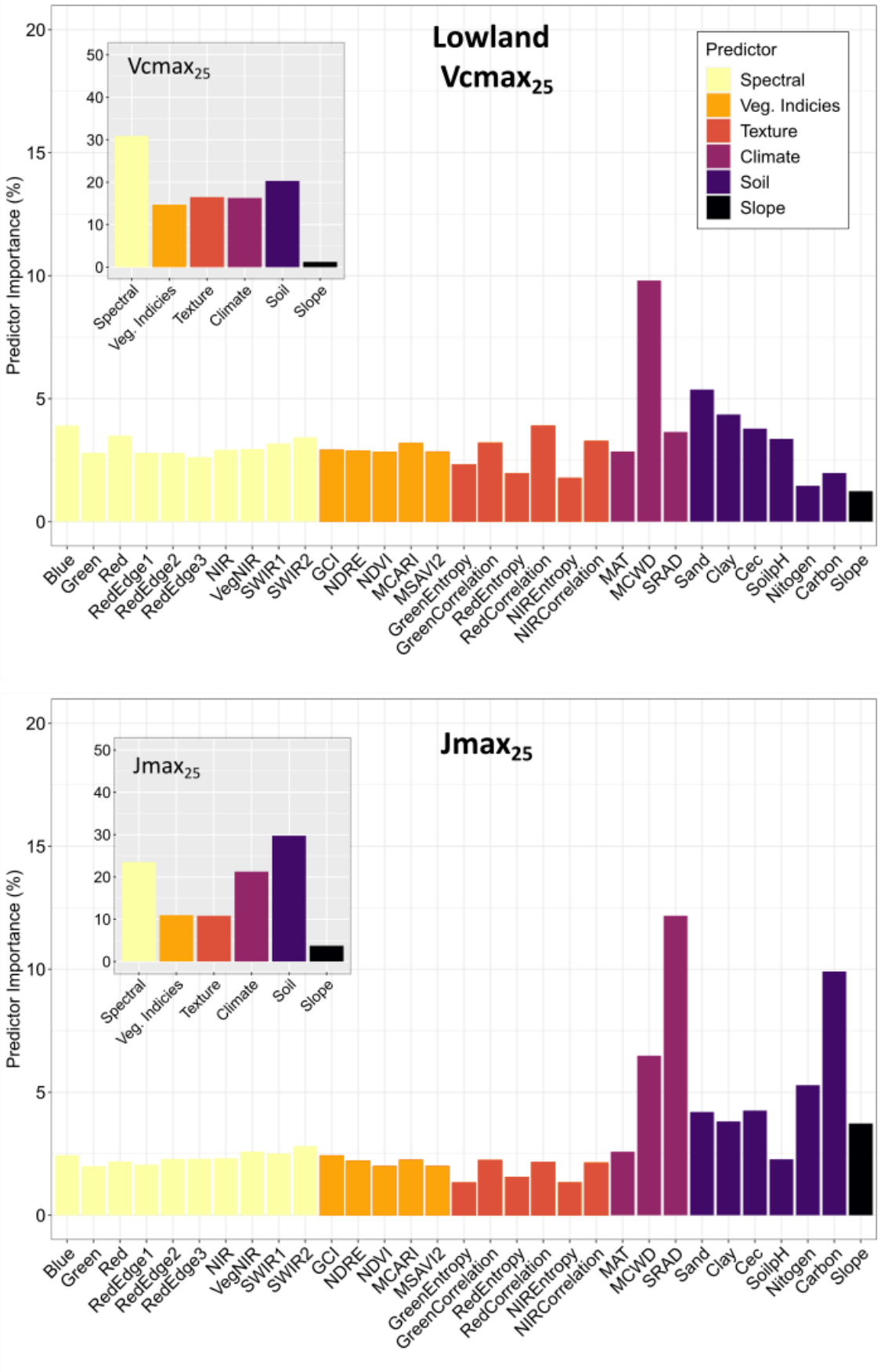


Figure 8.4: Percentage importance of each spectral, texture, climatic and soil variable for predicting lowland $Vcmax_{25}$ and $Jmax_{25}$. *Inset*: Total importance for each predictor group, represented as a percentage.

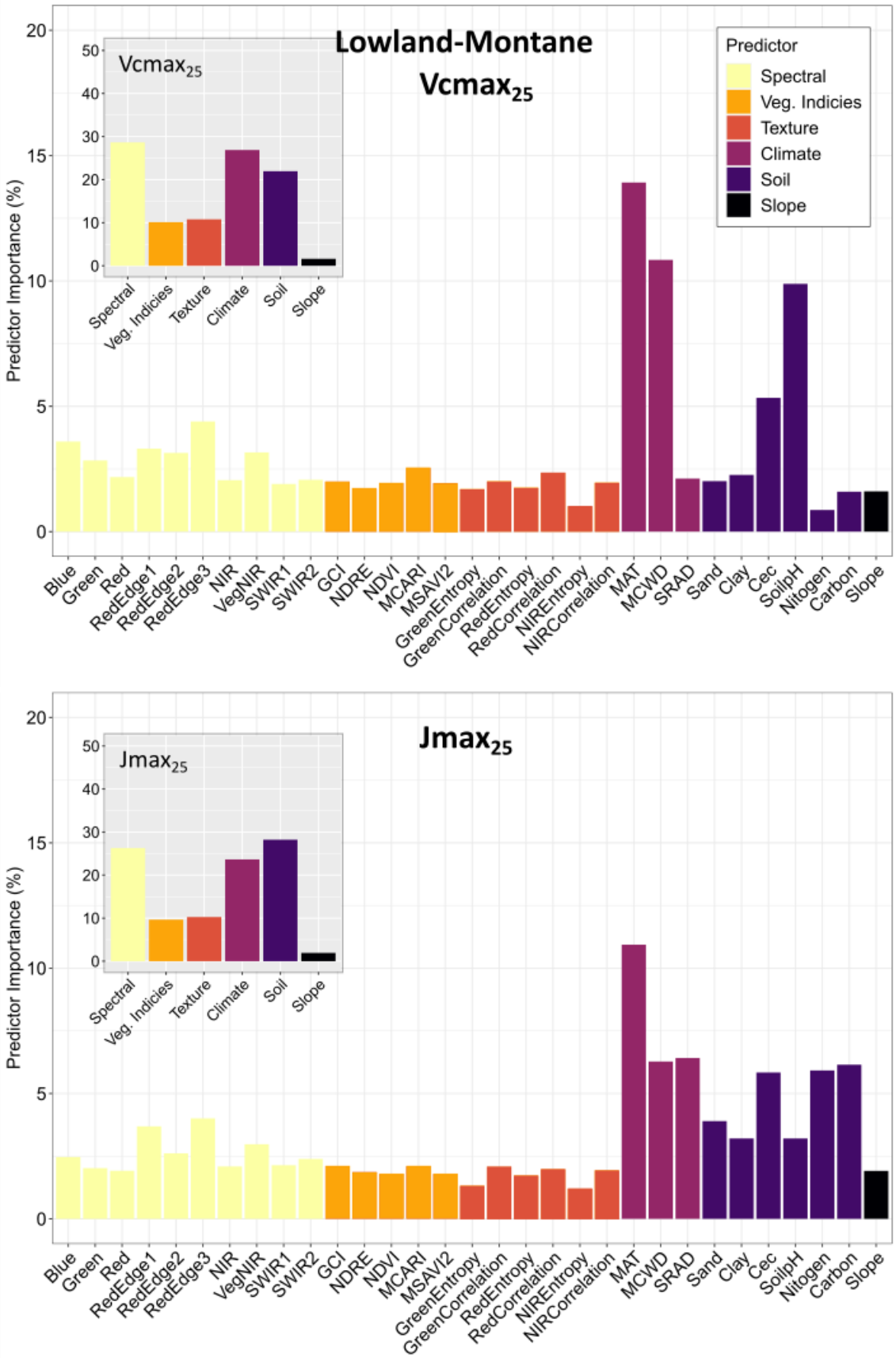


Figure 8.5: Percentage importance of each spectral, texture, climatic and soil variable for predicting lowland-montane $Vcmax_{25}$ and $Jmax_{25}$. *Inset*: Total importance for each predictor group, represented as a percentage.

In both the lowland only, and lowland-montane $Vcmax_{25}$ models, combined spectral variables were more important for prediction accuracy than environmental data (**Figure 8.4 and 8.5**). In the lowland only $Vcmax_{25}$ model, Sentinel bands contributed 31% to prediction performance, compared to climate variables (16%) and soil data (21%). Individual spectral bands and vegetation indices contributed similarly to prediction accuracy (around 4%). The predictor with the largest single importance for predicting lowland $Vcmax_{25}$ was MCWD, followed by %sand and %clay (**Figure 8.4**).

Across both $Jmax_{25}$ models (**Figure 8.4 and 8.5**), soil information (29%) was more important than spectral (23%) or climate data (21%) at predicting $Jmax_{25}$. In the lowland only $Jmax_{25}$ model (**Figure 8.4**), the single most important predictors were solar radiation (12%) followed by soil organic carbon content (10%), MCWD (6.5%) and nitrogen content (5.5%) In the lowland-montane models, MAT was the single most important predictor across both models (**Figure 8.5**).

None of the models found vegetation indices were more useful for predicting photosynthetic capacity than single Sentinel bands. Contributions from texture metrics were similar across all four models, with correlation contributing more to prediction accuracy than entropy.

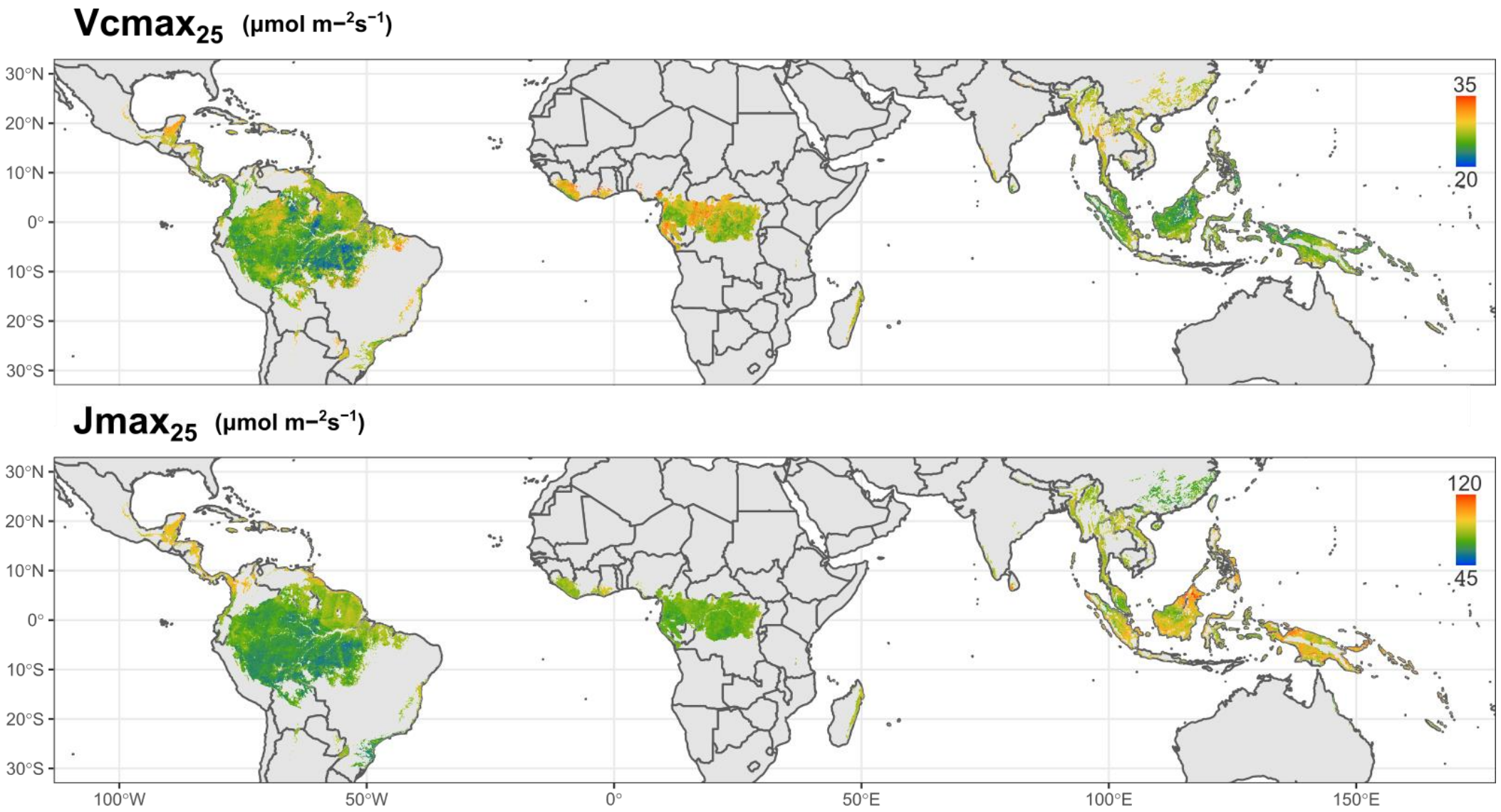


Figure 8.6: Pan-tropical maps of lowland $Vcmax_{25}$ and $Jmax_{25}$. Created in Google Earth Engine using field-collected trait values, Senitnel-2A imagery, environmental data and random forest regression. Original resolution at 10m. Resampled to 10km for display purposes. Only pixels located within the tropical forest biome, with >50% tree cover and <1000 m.a.s.l are classified.

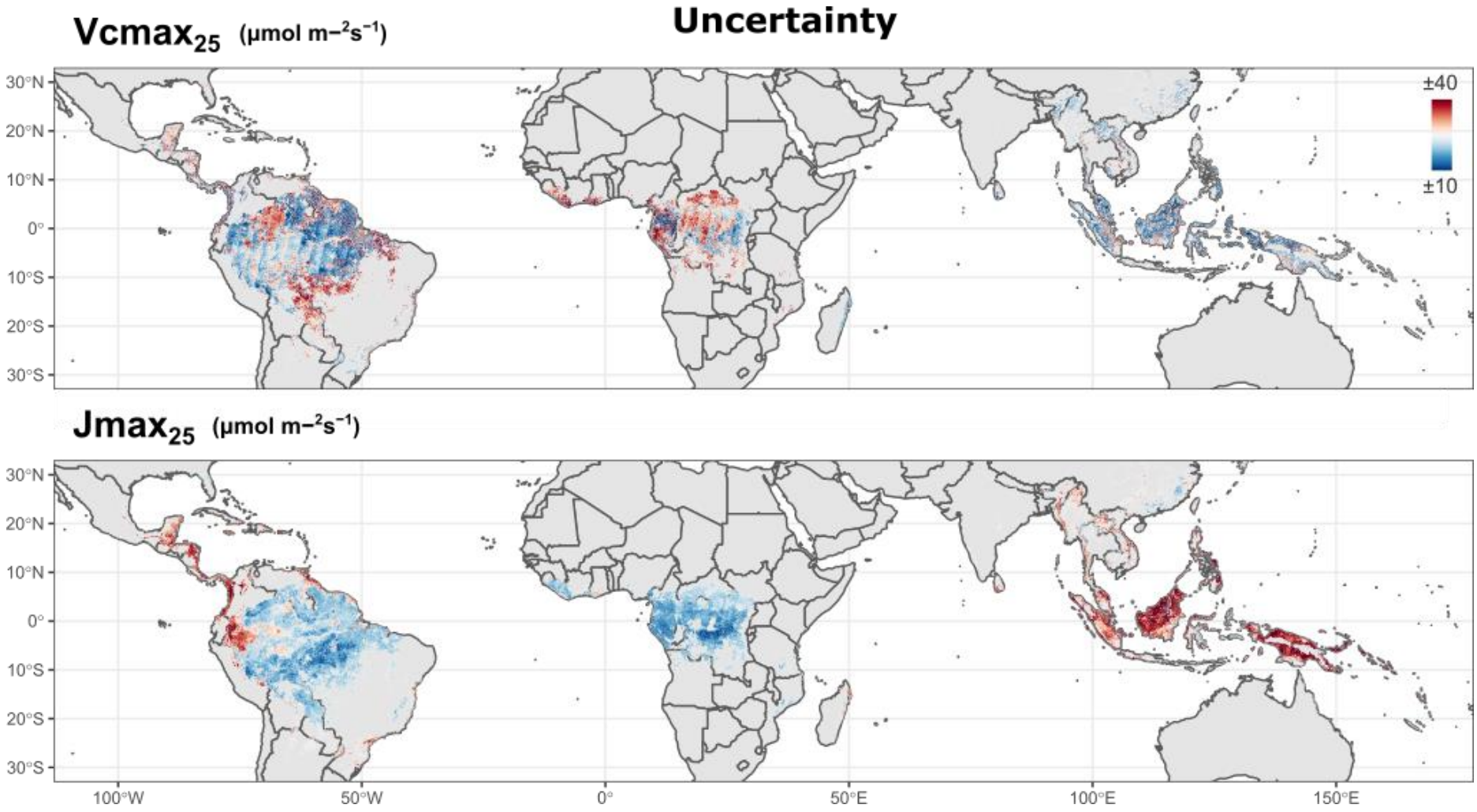


Figure 8.7: Uncertainty maps for $Vcmax_{25}$ and $Jmax_{25}$. Created in Google Earth Engine by calculating the standard deviation all random forest model tree predictions for each 10m pixel.

Our maps of lowland photosynthetic capacity show significant regional variation in $Vcmax_{25}$ and $Jmax_{25}$ (**Figure 8.6**). The highest $Vcmax_{25}$ values were displayed in African tropical forests, with particularly high $Vcmax_{25}$ (>35 μmol/m^{-2} s^{-1}) predicted across West Equatorial Africa, and the inland forests of West Africa. South America and Malaysia-Indonesia Archipelago displayed lower rates of $Vcmax_{25}$, with the lowest $Vcmax_{25}$ (<20 μmol/m^{-2} s^{-1}) predicted in Southern Amazonia and Borneo. Outside of the African continent, the highest $Vcmax_{25}$ values were predicted across the Yucatan Peninsula (>30 μmol/m^{-2} s^{-1}), the upper Rio Negro basin of northwest Amazonia (27-32 μmol/m^{-2} s^{-1}) and along the Cerrado-Amazonia transition (CAT).

Pan-tropical $Vcmax_{25}$ and $Jmax_{25}$ patterns demonstrated little correlation, indicating sensitivity to different environmental controls. Globally, the highest tropical forest $Jmax_{25}$ values were predicted across Malaysia-Indonesia (>115 μmol/m^{-2} s^{-1}). $Jmax_{25}$ values were uniform across the Congo Basin and West Africa (72-76 μmol/m^{-2} s^{-1}), $Jmax_{25}$ rates were slightly lower across the majority of South America (50-70 μmol), except along a strip of northwest coastal forest, which exhibited high $Jmax_{25}$ values (90-100 μmol/m^{-2} s^{-1}). $Jmax_{25}$ values were similarly high across Central America (85-95 μmol/m^{-2} s^{-1}).

In **Figure 8.8** we illustrated predicted maps for the two elevation gradients in our dataset. $Vcmax_{25}$ is predicted to have strong $Vcmax_{25}$-elevation relationships along altitudinal gradients in Peru and Australia. From the Peruvian Amazonian lowlands to the Andes treeline at 3500 m.a.s.l, predicted $Vcmax_{25}$ increased from ~20 to >50 μmol/m^{-2} s^{-1}. In Australia, forest was predicted to have higher lowland values (~25 μmol/m^{-2} s^{-1}) and attain greatest $Vcmax_{25}$ (~40 μmol/m^{-2} s^{-1}) at around 1000m.

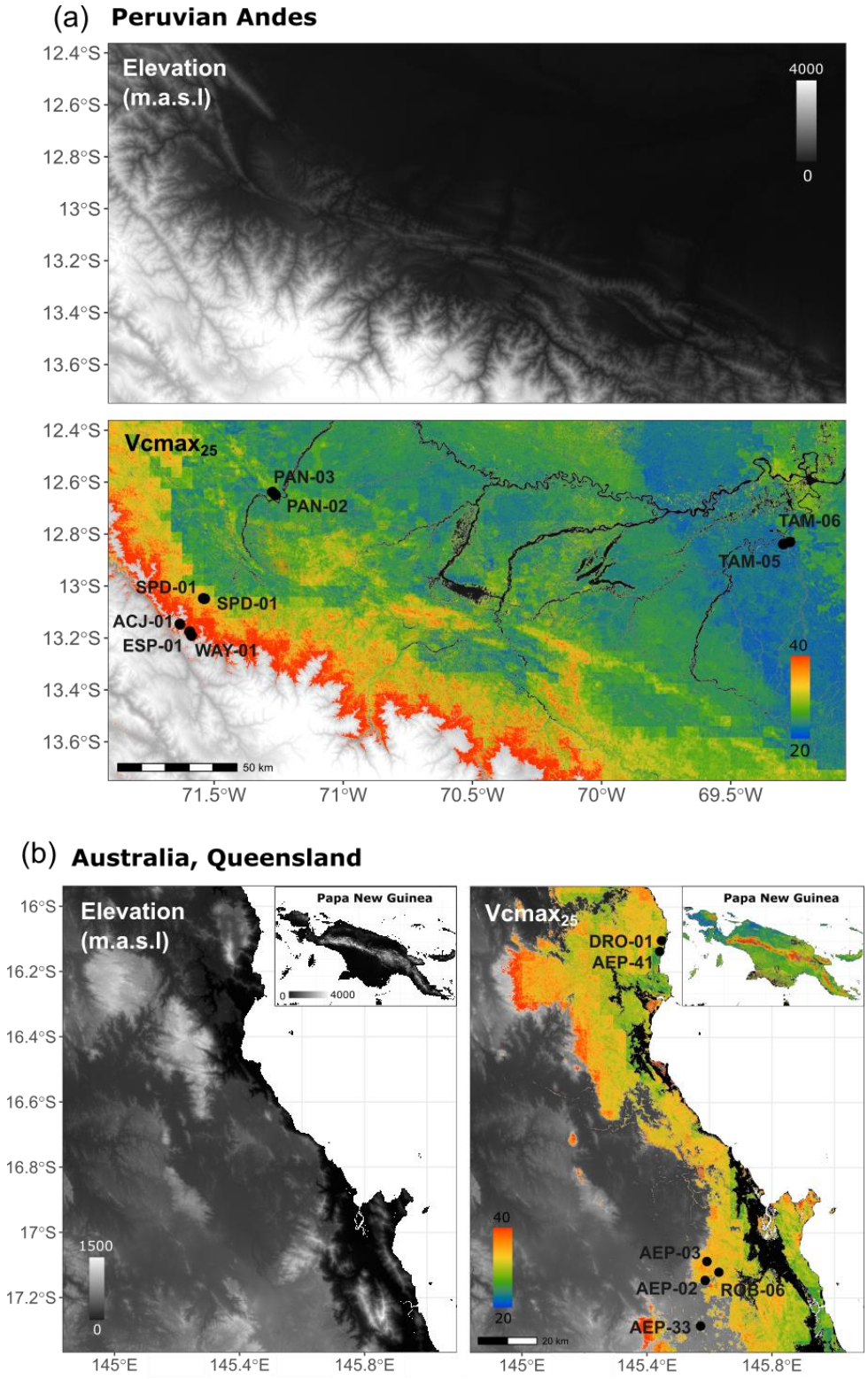


Figure 8.8*:* Maps of elevation and lowland-montane Vcmax$_{25}$ (a) an elevational gradient in Peru from the Amazon lowlands (200–225 m.a.s.l.), across the Pantiacolla front range of the Andes (600–900 m.a.s.l.), to the Kosñipata Valley (1,500–3,500 m). (b) A coastal-inland elevation gradient in Australia, Queensland with an insert for nearby Papua New Guinea. Created in Google Earth Engine using field-collected trait values. Senitnel-2A imagery, environmental data and random forest regression. Spatial resolution 10m.

8.7. Discussion

To the best of our knowledge, this is the first research that uses Sentinel-2 imagery to create pan-tropical maps of $Vcmax_{25}$ and $Jmax_{25}$. Scaling from the leaf, to the crown, to the pixel level, we have shown that accurate (r^2=0.31-0.62) predictions of photosynthetic capacity can be achieved at 10x10m, using field-measured photosynthesis values of the largest most abundant canopy species, combined with geolocated stem and crown positions, Sentinel-2 imagery and environmental data layers.

The 'one-point' method

Testing the 'one-point' method at tropical forest sites across three continents, produced strong relationships (r = 0.76-96) between $Vcmax_{25}$ and $Jmax_{25}$ values derived from traditional *A*-Ci-curves ('the gold standard'), and those estimated using 'one-point' A_{sat}, A_{max} and R_{dark} measurements (**Figure S8.1**). As *A*-Ci-curves can take up to an hour to generate, compared to <2 minutes for a one-point measurement, this result is promising for expanding current datasets of photosynthetic capacity across the tropics and accelerating future data collection for model parametrization.

Field-sampled photosynthetic capacity distributions

Field-sampled $Vcmax_{25}$ and $Jmax_{25}$ values differed significantly across tropical continents and along environmental gradients (**Figure 8.2**). $Vcmax_{25}$ was lowest in the humid lowland plots of Malaysian Borneo and Amazonia, and highest in the Australian and African plots. These patterns may reflect differences in soil age and fertility, accentuated by climate. According to Albrecht's conceptual model of soil degradation, soil P and N should decline along a gradient from low to high rainfall and from young to older soils (Huston, 2012). Empirical studies have shown tight correlations between leaf nitrogen content and photosynthetic capacity (e.g. Field and Mooney 1986; Reich *et al*. 1998; Medlyn *et al*. 1999; Walker *et al*. 2014) with a constant relationship between leaf nitrogen and $Vcmax_{25}$ utilized in many ecosystem models (Haxeltine and Prentice, 1996; Bonan *et al*. 2003; Kattge *et al*. 2009). In phosphorus-limited environments, such as many parts of the tropics, increasing leaf P has been found to substantially increase the sensitivity of $Vcmax_{25}$ to leaf nitrogen (Walker *et al*. 2014), leading to stronger soil P- $Vcmax_{25}$ than soil

N- $Vcmax_{25}$ relationships (Meir, 2001; Meir *et al.* 2007; Domingues *et al*. 2010; Bahar *et al*. 2014). Thus, our patterns of photosynthetic capacity adhere to Albrecht's model, with lowest $Vcmax_{25}$ found on the oldest, most weathered soils across lowland Amazonia and Borneo (MAP: >2000mm), and higher photosynthetic capacity displayed in tropical forest growing in comparatively younger soils with a lower precipitation-evapotranspiration balance (MAP: <2000mm).

Across environmental gradients, elevation and forest disturbance were found to have the greatest effect on photosynthetic capacity. Ascending the Peruvian lowland-montane gradient, $Vcmax_{25}$ increased from 21 ± 6 to 91 ± 15 μmol/m^{-2} s^{-1}. Similar results were found by Bahar *et al*. (2016: 6–96 μmol/m^{-2} s^{-1}) and van de Weg *et al*. (2011) along the same gradient. Enhanced $Vcmax_{25}$ has also been found in montane forest in Rwanda (Dusenge *et al*. 2005; Vårhammar *et al.* 2015) and along latitudinal temperature gradients (Kattge and Knorr *et al*. 2007; Crous *et al*. 2018). At high altitude, it has been hypothesised that enhanced photosynthetic capacity helps negate the inhibitory effects of low MAT on photosynthetic rate, facilitated by larger N investments into photosynthetic enzymes to compensate for slower enzyme activity (Bahar *et al.* 2017). A weaker elevation effect was displayed across the Australian plots (22 ± 9 to 36 ± 10 μmol/m^{-2} s^{-1}), likely due to the lower elevation, and the disrupting effect of seasonality and differing forest soils: AEP-02 is a seasonally dry plot, AEP-41 and DRO-01 are moderately seasonal, while AEP-33 is the least seasonal plot. Most of the plots are located on weathered soils while DRO-01 is located on unweathered basalt soils.

Forest disturbance also had a strong impact on $Vcmax_{25}$ and $Jmax_{25}$. Increased photosynthetic capacity was observed in logged forest relative to nearby old-growth forest, in both the Brazilian and Malaysian Borneo plots, likely driven by the increased presence of pioneer or other fast-growing species. Regenerating in high-light environments, fast-growing species tend to invest more in leaf traits that maximise photosynthetic capacity, compared to climax species regenerating in the shade of the forest canopy that tend to invest more resources in light-harvesting (Raaimakers *et al.* 1995; Valladares and Niinemets, 2008). In Indonesian forest, Hölscher *et al*. (2005) found photosynthetic capacity was 2.3 times higher in early-successional *Macaranga* species than old grown *Dipterocarp* species. Little variation in photosynthetic capacity was

observed along the precipitation gradient in Ghana. While $Vcmax_{25}$ was significantly lower in the defaunation plot in Gabon, lower $Vcmax_{25}$ is likely due to the sandy, highly weathered soils of Mondah compared to Lopé.

Some of our $Vcmax_{25}$ values were on the low end of ranges reported for lowland rainforests, but were within the ranges reported for $Jmax_{25}$ (**Table S8.6**). The wide range of $Vcmax_{25}$ and $Jmax_{25}$ values measured within and across our sites, and wider pan-tropical studies, highlights the importance and challenge of building comprehensive datasets on inter-continental photosynthetic capacity distributions. These datasets are of pivotal importance for elucidating the global, regional and local controls on photosynthetic capacity, and calibrating and validating TBMs.

Model predictions

Pan-tropical models of lowland photosynthesis capacity were able to explain 10x10m variation in $Vcmax_{25}$ and $Jmax_{25}$ at r^2=0.38 and 0.57 when all pixels were involved in training and validation, and r^2=0.16 and 0.30 when a spatial fold method of validation was used (**Table 8.3**). This difference in accuracy is related to the non-spatial *vs*. spatial method of validation. Using a random forest model and similar multispectral and environmental variables, Ploton *et al*. (2020) showed how standard cross-validation could produce a model with apparently high predictive power for aboveground biomass, but reveal quasi-null predictive power when spatial autocorrelation effects were account for. According to Ploton *et al*. (2020), using a standard cross-validation that treats spatial pixels as independent is flawed, as dependence structures exist across the data (i.e. observations close to each other tend to have similar characteristics), leading to an overoptimistic assessment of model predictive power. Therefore Ploton *et al*. (2020), recommend using a spatial cross-validation approach that ensures an observation point cannot serve as an independent validation of a nearby training data point. However, according to Wadoux *et al*. (2021), while standard cross-validation may be deficient in the case of a clustered dataset (such as ecological plots), standard cross-validation is less deficient than spatial cross-validation. In spatial cross-validation, geographically-structured ecological and environmental variables might be unseen in the calibration

folds, forcing the model to predict outside the feature space covered by covariates and resulting in overly-pessimistic validation results.

We include both methods of cross-validation in our analysis and suggest that real model accuracy lies somewhere in the middle, and potentially closer to the standard cross-validation results. Despite the mismatch between the environmental data layers (90m-4000m) and spatial resolution of the trait data (10m), we do not believe that our dataset suffers from severe SAC effects. This is due to the large range of $Vcmax_{25}$ and $Jmax_{25}$ trait values recorded within each plot (often spanning the entire range of values recorded pan-tropically), which would have hindered model performance, rather than leading to enhanced model results. In addition, the inclusion of environmentally unique plots in our analysis (such as the upper montane ACJ-01, or the Kogyae savannah plots) would have unduly penalised a spatial fold validation. When ACJ-01 appeared in the validation dataset, the model, bounded by its calibration values, would have been physically unable to predict into such high trait space. When the Kogyae spatial fold appeared in the validation dataset, the model would have lacked covariates pertaining to forest-transition conditions. Thus, the strength of a random forest model is also its weakness. Random forest is a flexible, non-parametric model that can handle non-linear trait-predictor relationships and interactions between predictors (such as increased $Vcmax_{25}$ with soil N in the presence of soil P). However, because of its empirical nature, a random forest model relies heavily on the target values and predictors present in its training dataset, and cannot predict into novel environmental or trait space.

Pan-tropical lowland-montane models captured more variation in $Vcmax_{25}$and $Jmax_{25}$ values than lowland models (r^2 = 0.67 and 0.73 respectively), reflecting the strong relationship between photosynthetic capacity and MAT. The most accurate local models were found for $Vcmax_{25}$ and $Jmax_{25}$ across the Peruvian elevation gradient (r^2 = 0.72 and 0.73). In contrast, models poorly explained local variation in photosynthetic capacity at lowland sites (r^2 = 0.06-0.32), with the strongest models arising from the forest disturbance gradients in Brazil-ST and Malaysia. Poor local prediction indicates that models struggled to estimate photosynthetic capacity within plots, with trait-predictor relationships arising from differences between plot means. Similar results have been found using airborne hyperspectral data over Borneo (Swinfield *et al*. 2019; **Chapter 7**).

Poor plot-level prediction may be due to a number of reasons. The mismatch between the environmental data layers (90m-4000m) and spatial resolution of the trait data (10m), may have failed to capture the presence of micro-climates and micro-soil variations influencing trait distributions at the 10x10m pixel level. Sampling 3-5 leaves per tree may not represent the full range of intra-crown photosynthetic capacity (Katajima *et al*. 1996, 1997; Carswell *et al*. 2000), or mismatches could exist between the tree crown maps and the Sentinel data, leading to inaccuracies in the spatial distribution of pixel trait values across a plot (Gascon *et al*. 2017). While some of these problems could be minimised with more finely resolved environmental data, or better geolocated satellite imagery, the remote sensing products used in this study represent the highest-resolution, most accurate datasets available. Therefore, errors associated with remote sensing base layers will only reduce with increased investment into space-based measurements and associated technology.

Predictor importance

The strongest predictor for lowland $Vcmax_{25}$ was MCWD (**Figure 8.4**). In drier environments, high $Vcmax_{25}$ has been found to be a strategy for minimising water loss, with increased investment in Rubisco facilitating reduced stomatal conductance (Wright *et al*. 2003; Prentice *et al*. 2014). The second and third strongest predictors for $Vcmax_{25}$ were % soil sand and % soil clay. As the model lacked information on soil phosphorus, soil sand and clay concentrations may act as reasonable proxies. 'White-sand' Amazonian sites are chronically low in soil phosphorus (Herrera *et al*. 1978; Quesada *et al*. 2010), while clay forms a more complicated relationship, with low to moderate clay content helping retain soil nutrients, and high clay concentrations binding to phosphorus cations, reducing labile availability.

The strongest predictor of lowland $Jmax_{25}$ was solar radiation (**Figure 8.4**), which may be associated with $Jmax_{25}$ being a light-limited process. Across all predictor groups, $Jmax_{25}$ showed greater sensitivity to soil conditions than Vcmax, similarly reported by Bahar *et al*. (2016) who found $Jmax_{25}$ exhibited stronger relationships with soil P and soil N than $Vcmax_{25}$. For $Vcmax_{25}$, spectral imagery was the most important predictor group. While abiotic data can shed light on the mechanistic relationships between traits and

environmental variables, spectral data acts as an 'indicative' predictor. Spectral imagery has been used to estimate $Vcmax_{25}$ through relationships with visible and NIR spectral bands (Serbin *et al*. 2015; **Chapter 7**), while a satellite-based hyperspectral index of vegetation chlorophyll has been used to map $Vcmax_{25}$ at a global scale (Alton, 2008).

All Sentinel-2 spectral bands contributed to the prediction performance of $Vcmax_{25}$ and $Jmax_{25}$ models, suggesting that hyperspectral satellite imagery, combined with our environmental predictors may provide the best near-term improvement to global vegetation models. In contrast to results found for Landsat and Modis studies (Houborg *et al*. 2013; Croft *et al*. 2016), vegetation indices were less important than the raw spectral bands across all models. Compared to related ecological mapping studies in tropical forests (Laurin *et al.* 2016; Aguirre-Gutiérrez *et al*. 2020), we found the contribution of texture metrics was lower than expected. However, like Laurin *et al*. (2016), we found that correlation was more important than entropy for characterising texture and capturing ecological diversity.

Pan-tropical maps of $Vcmax_{25}$ and $Jmax_{25}$

Pan-tropical maps of lowland photosynthetic capacity displayed large inter-continental differences in $Vcmax_{25}$ and $Jmax_{25}$ controlled by climate (**Figure 8.6**). $Vcmax_{25}$ was highest in Western Equatorial Africa (excluding a ferrasol intrusion: see below) and mainland south-east Asia. $Vcmax_{25}$ was lowest across most of the neotropics and the Malay Archipelago. Western Equatorial Africa is generally much drier than most of the neotropics and south-east Asia and humid evergreen forest can only be maintained in this region due to reduced water demand associated with extreme cloudiness (Philippon *et al*. 2019). Only 3% of total annual rainfall falls during the long dry season from June to September (Balas *et al*. 2007), with high $Vcmax_{25}$ likely required as a strategy to reduce water loss during these months. In contrast, parts of the neotropics and insular Southeast Asia have some of the highest year-round precipitation rates on the planet. Regionally, relationships with MCWD could be observed across MCWD gradients spanning Liberia, the Yucatan peninsula, and the Cerrado-Amazonia Transition (CAT).

Intra-continental patterns of $Vcmax_{25}$ reflected the distribution of soil nutrients. In Western Equatorial Africa, comparatively low $Vcmax_{25}$ (23-26 μmol/m^{-2} s^{-1}) was predicted over a clay-rich ferrosol intrusion spanning northern Gabon, Equatorial Guinea and southern Cameroon. In contrast, high $Vcmax_{25}$ (>35 μmol/m^{-2} s^{-1}) was predicted in interfluvial basins of the Cuvette Centrale depression corresponding to swamp forest overlying thick peat deposits (Dargie *et al.* 2017). In northwest Amazonia, high $Vcmax_{25}$ was shown over an area of Arenosols and Podzols in the upper Rio Negro basin (Quesada *et al*. 2008), with low $Vcmax_{25}$ found over a band of 'white sand' running through the Tambopata plots (**Figure 8.8**).

The interrelationship of $Vcmax_2$ and environmental variables at different scales, helps reconcile the climate *vs* nutrient limitation debate. While nutrient (especially N) limitation has classically been used as the main control on $Vcmax_{25}$ (Rogers *et al*. 2013; Walker *et al*. 2017), recent global analyses have found that the majority of global $Vcmax_{25}$ variation can be explained by water and light availability, with only a small contribution from soil fertility (Prentice *et al*. 2014; Ali *et al*. 2015; Smith *et al*. 2019). Our results support these recent studies, and show that, across large spatial scales, MCWD is the principle control on pan-tropical $Vcmax_{25}$, with soil fertility mediating regional-local distributions.

Pan-tropical patterns of $Jmax_{25}$ were mainly controlled by solar radiation. $Jmax_{25}$ was greatest in areas of high solar radiation, such as northern Borneo and parts of the Malay peninsula, the Guianas, and central America. However, $Jmax_{25}$ showed less sensitivity to typically cloudy regions, such as Gabon and northwest Amazonia. A lack of $Jmax_{25}$ sensitivity in cloudy regions implies that lower rates of $Jmax_{25}$ may be bounded by $Vcmax_{25}$, with plants investing strongly in $Jmax_{25}$ in high insolation, but still requiring a minimum rate of RuPB regeneration to support RuBiSCO activity, even under low light levels (Smith *et al*. 2019). Supporting this hypothesis, the lowest $Jmax_{25}$ values were corelated with patterns of low $Vcmax_{25}$ in south-east Amazonia (<60 μmol/m^{-2} s^{-1}).

Montane forest displayed a tight relationship between $Vcmax_{25}$ and MAT (**Figure 8.8**). Across the Peruvian elevation gradient, this relationship was not linear. At 950m, there was a sharp increase in $Vcmax_{25}$ from 25 to >30 μmol/m^{-2} s^{-1}, and another jump at

3000m to >50 μmol/m^{-2} s^{-1}. In this region, 3000m represents the top of the cloud line (Halleday *et al*. 2012), suggesting that for the highest montane forest, rapid increases in $Vcmax_{25}$ can be achieved when $Vcmax_{25}$ is no longer limited by $Jmax_{25}$. While elevation was the principle large-scale control of $Vcmax_{25}$, at local scales, spectral information became the dominant control over trait variation, with micro-patterning of $Vcmax_{25}$ associated with local topography and canopy structure.

Uncertainty maps showed a greater range of predictions at higher values of $Vcmax_{25}$ and $Jmax_{25}$, suggesting heteroskedasticity in the model (**Figure 8.7**). In addition, uncertainty was highest in areas forest unrepresented by our 35 plots, such as *varillo* vegetation in northwest Amazonia, the CAT, central America and mainland south-east Asia. Our maps of $Vcmax_{25}$ and $Jmax_{25}$ will take a great deal of effort to validate. However, uncertainty maps can help to direct this effort towards the most under-sampled forests, to increase representation in these areas and improve model consensus.

8.8. Conclusion

Spatially explicit information on photosynthetic capacity, especially across the tropics, is required to generate robust projections of regional and global environmental change. In addition, $Vcmax_{25}$ patterns have been found to be indicative of drought (Xu and Baldocchi, 2003), insect and pathogen damage (Dungan *et al*. 2007) and air pollution (Ainsworth *et al*. 2014).

This study investigates the relationships between photosynthetic capacity and spectral, textual and environmental predictors at local and pan-tropical scales. Using random forest regression, we demonstrate a new approach for mapping photosynthetic capacity ($Vcmax_{25}$ and $Jmax_{25}$) at high spatial resolution (10x10m) across the tropics. We find that, at intercontinental scales, $Vcmax_{25}$ and $Jmax_{25}$ are mainly controlled by MCWD and solar radiation. At the regional level, trait distributions are shaped by soil type and elevation. At local levels, trait patterns reflect spectral variation arising from topography and canopy structure.

Machine learning approaches are data hungry and benefit from the inclusion of new data, especially in unrepresented trait or covariate space. Building up a global database of geolocated photosynthetic capacity measurements will strengthen these kind of model analyses and allow changes in forest ecosystem functioning to be tracked across space and time, in one of the greatest eras of environmental change.

8.9. References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A. and Hegewisch, K.C. (2018) TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific data,* 5 (1), pp. 1-12.

Aguirre-Gutiérrez, J., Rifai, S., Shenkin, A., Oliveras, I., Bentley, L.P., Svátek, M., Girardin, C.A.J., Both, S., Riutta, T. and Berenguer, E. (2021) Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data. *Remote Sensing of Environment,* 252, pp. 112122.

Ainsworth, E.A., Serbin, S.P., Skoneczka, J.A. and Townsend, P.A. (2014) Using leaf optical properties to detect ozone effects on foliar biochemistry. *Photosynthesis Research,* 119 (1), pp. 65-76.

Ali, A.A., Xu, C., Rogers, A., Fisher, R.A., Wullschleger, S.D., Massoud, E.C., Vrugt, J.A., Muss, J.D., McDowell, N.G. and Fisher, J.B. (2016) A global scale mechanistic model of photosynthetic capacity (LUNA V1. 0). *Geoscientific Model Development,* 9 (2), pp. 587-606.

Alton, P.B. (2018) Decadal trends in photosynthetic capacity and leaf area index inferred from satellite remote sensing for global vegetation types. *Agricultural and Forest Meteorology,* 250, pp. 361-375.

Asner, G.P. (2001) Cloud cover in Landsat observations of the Brazilian Amazon. *International Journal of Remote Sensing,* 22 (18), pp. 3855-3862.

Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E. and Vaughn, N. (2016) Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proceedings of the National Academy of Sciences,* 113 (28), pp. E4043-E4051.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn, N.R. and Llactayo, W. (2017) Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science,* 355 (6323), pp. 385-389.

Asner, G.P. and Martin, R.E. (2008) Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment,* 112 (10), pp. 3958-3970.

Asner, G.P., Martin, R.E., Anderson, C.B. and Knapp, D.E. (2015) Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sensing of Environment,* 158, pp. 15-27.

Asner, G.P., Martin, R.E., Tupayachi, R., Anderson, C.B., Sinca, F., Carranza-Jiménez, L. and Martinez, P. (2014) Amazonian functional diversity from forest canopy chemical assembly. *Proceedings of the National Academy of Sciences,* 111 (15), pp. 5604-5609.

Bahar, N.H.A., Ishida, F.Y., Weerasinghe, L.K., Guerrieri, R., O'Sullivan, O.,S., Bloomfield, K.J., Asner, G.P., Martin, R.E., Lloyd, J. and Malhi, Y. (2017) Leaf-level photosynthetic capacity in lowland Amazonian and high-elevation Andean tropical moist forests of Peru. *New Phytologist,* 214 (3), pp. 1002-1018.

Balas, N., Nicholson, S.E. and Klotter, D. (2007) The relationship of rainfall variability in West Central Africa to sea-surface temperature fluctuations. *International Journal of Climatology: A Journal of the Royal Meteorological Society,* 27 (10), pp. 1335-1349.

Barnes, M.L., Breshears, D.D., Law, D.J., Van Leeuwen, W.,J.D., Monson, R.K., Fojtik, A.C., Barron-Gafford, G. and Moore, D.J.P. (2017) Beyond greenness: detecting temporal changes in photosynthetic capacity with hyperspectral reflectance data. *PLoS One,* 12 (12), pp. e0189539.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M.A., Baldocchi, D. and Bonan, G.B. (2010) Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science,* 329 (5993), pp. 834-838.

Berenguer, E., Malhi, Y., Brando, P., Cardoso Nunes Cordeiro, A., Ferreira, J., França, F., Chesini Rossi, L., Maria Moraes de Seixas, Marina and Barlow, J. (2018) Tree growth and stem carbon accumulation in human-modified Amazonian forests following drought and fire. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 373 (1760), pp. 20170308.

Bloomfield, K.J., Farquhar, G.D. and Lloyd, J. (2014) Photosynthesis–nitrogen relationships in tropical forest tree species as affected by soil phosphorus availability: a controlled environment study. *Functional plant biology,* 41 (8), pp. 820-832.

Bonan, G.B. (2008) Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science,* 320 (5882), pp. 1444-1449.

Bonan, G.B., Lawrence, P.J., Oleson, K.W., Levis, S., Jung, M., Reichstein, M., Lawrence, D.M. and Swenson, S.C. (2011) Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *Journal of Geophysical Research: Biogeosciences,* 116.

Bonan, G.B., Levis, S., Sitch, S., Vertenstein, M. and Oleson, K.W. (2003) A dynamic global vegetation model for use with climate models: concepts and description of simulated vegetation dynamics. *Global Change Biology,* 9 (11), pp. 1543-1566.

Both, S., Riutta, T., Paine, C.T., Elias, D.M.O., Cruz, R.S., Jain, A., Johnson, D., Kritzler, U.H., Kuntz, M. and Majalap-Lee, N. (2019) Logging and soil nutrients independently explain plant trait expression in tropical forests. *New Phytologist,* 221 (4), pp. 1853-1865.

Breiman, L. (2001) Random forests. *Machine Learning,* 45 (1), pp. 5-32.

Burnett, A.C., Davidson, K.J., Serbin, S.P. and Rogers, A. (2019) The "one-point method" for estimating maximum carboxylation capacity of photosynthesis: a cautionary tale. *Plant, Cell & Environment,* 42 (8), pp. 2472-2481.

Carswell, F.E., Meir, P., Wandelli, E.V., Bonates, L., Kruijt, B., Barbosa, E.M., Nobre, A.D., Grace, J. and Jarvis, P.G. (2000) Photosynthetic capacity in a central Amazonian rain forest. *Tree physiology,* 20 (3), pp. 179-186.

Cleveland, C.C., Taylor, P., Chadwick, K.D., Dahlin, K., Doughty, C.E., Malhi, Y., Smith, W.K., Sullivan, B.W., Wieder, W.R. and Townsend, A.R. (2015) A comparison of plot-based satellite and Earth system model estimates of tropical forest net primary production. *Global Biogeochemical Cycles,* 29 (5), pp. 626-644.

Croft, H., Chen, J.M., Luo, X., Bartlett, P., Chen, B. and Staebler, R.M. (2017) Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global Change Biology,* 23 (9), pp. 3513-3524.

Crous, K.Y., Drake, J.E., Aspinwall, M.J., Sharwood, R.E., Tjoelker, M.G. and Ghannoum, O. (2018) Photosynthetic capacity and leaf nitrogen decline along a controlled climate gradient in provenances of two widely distributed Eucalyptus species. *Global Change Biology,* 24 (10), pp. 4626-4644.

Curran, P.J. (1989) Remote sensing of foliar chemistry. *Remote Sensing of Environment,* 30 (3), pp. 271-278.

Dargie, G.C., Lewis, S.L., Lawson, I.T., Mitchard, E.T.A., Page, S.E., Bocko, Y.E. and Ifo, S.A. (2017) Age, extent and carbon storage of the central Congo Basin peatland complex. *Nature,* 542 (7639), pp. 86-90.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., De Colstoun, E.B. and McMurtrey Iii, J. (2000) Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment,* 74 (2), pp. 229-239.
De Kauwe, M.,G., Lin, Y., Wright, I.J., Medlyn, B.E., Crous, K.Y., Ellsworth, D.S., Maire, V., Prentice, I.C., Atkin, O.K. and Rogers, A. (2016) A test of the 'one-point method'for estimating maximum carboxylation capacity from field-measured, light-saturated photosynthesis. *New Phytologist,* 210 (3), pp. 1130-1144.
de Sousa, L.,M., Poggio, L., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Riberio, E. and Rossiter, D. (2020) SoilGrids 2.0: producing quality-assessed soil information for the globe. *Soil Discuss.,* 1.
Dechant, B., Cuntz, M., Vohland, M., Schulz, E. and Doktor, D. (2017) Estimation of photosynthesis traits from leaf reflectance spectra: Correlation to nitrogen content as the dominant mechanism. *Remote Sensing of Environment,* 196, pp. 279-292.
Dillen, S.Y., de Beeck, M.O., Hufkens, K., Buonanduci, M. and Phillips, N.G. (2012) Seasonal patterns of foliar reflectance in relation to photosynthetic capacity and color index in two co-occurring tree species, Quercus rubra and Betula papyrifera. *Agricultural and Forest Meteorology,* 160, pp. 60-68.
Domingues, T.F., Martinelli, L.A. and Ehleringer, J.R. (2007) Ecophysiological traits of plant functional groups in forest and pasture ecosystems from eastern Amazonia, Brazil. *Plant Ecology,* 193 (1), pp. 101-112.
Domingues, T.F., Ishida, F.Y., Feldpausch, T.R., Grace, J., Meir, P., Saiz, G., Sene, O., Schrodt, F., Sonké, B. and Taedoumg, H. (2015) Biome-specific effects of nitrogen and phosphorus on the photosynthetic characteristics of trees at a forest-savanna boundary in Cameroon. *Oecologia,* 178 (3), pp. 659-672.
Domingues, T.F., Meir, P., Feldpausch, T.R., Saiz, G., Veenendaal, E.M., Schrodt, F., Bird, M., Djagbletey, G., Hien, F. and Compaore, H. (2010) Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. *Plant, Cell & Environment,* 33 (6), pp. 959-980.
Doughty, C.E., Asner, G.P. and Martin, R.E. (2011) Predicting tropical plant physiology from leaf and canopy spectroscopy. *Oecologia,* 165 (2), pp. 289-299.
Doughty, C.E., Santos-Andrade, P., Goldsmith, G.R., Blonder, B., Shenkin, A., Bentley, L.P., **Chavana-Bryant, C., Huaraca-Huasco, W., Díaz, S. and Salinas, N. (2017)** Can leaf spectroscopy predict leaf and forest traits along a Peruvian tropical forest elevation gradient? *Journal of Geophysical Research: Biogeosciences,* 122 (11), pp. 2952-2965.
Dungan, R.J., Turnbull, M.H. and Kelly, D. (2007) The carbon costs for host trees of a phloem-feeding herbivore. *Journal of Ecology,* 95 (4), pp. 603-613.
Dusenge, M.E., Wallin, G., Gårdesten, J., Niyonzima, F., Adolfsson, L., Nsabimana, D. and Uddling, J. (2015) Photosynthetic capacity of tropical montane tree species in relation to leaf nutrients, successional strategy and growth temperature. *Oecologia,* 177 (4), pp. 1183-1194.
Duursma, R.A. (2015) Plantecophys-an R package for analysing and modelling leaf gas exchange data. *PloS one,* 10 (11), pp. e0143346.
Farquhar, G.D., Von Caemmerer, S. and Berry, J.A. (2001) Models of photosynthesis. *Plant Physiology,* 125 (1), pp. 42-45.
Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., **Rodriguez, E. and Roth, L. (2007)** The shuttle radar topography mission. *Reviews of Geophysics,* 45 (2).
Field, C.H. (1986) The photosynthesis-nitrogen relationship in wild plants. *On the economy of form and function,* .

Foley, J.A., Costa, M.H., Delire, C., Ramankutty, N. and Snyder, P. (2003) Green surprise? How terrestrial ecosystems could affect earth's climate. *Frontiers in Ecology and the Environment,* 1 (1), pp. 38-44.
Friend, A.D. (2010) Terrestrial plant production and climate change. *Journal of experimental botany,* 61 (5), pp. 1293-1309.
Gao, Y., Yu, G., Yan, H., Zhu, X., Li, S., Wang, Q., Zhang, J., Wang, Y., Li, Y. and Zhao, L. (2014) A MODIS-based Photosynthetic Capacity Model to estimate gross primary production in Northern China and the Tibetan Plateau. *Remote Sensing of Environment,* 148, pp. 108-118.
Gara, T.W., Darvishzadeh, R., Skidmore, A.K., Wang, T. and Heurich, M. (2019) Accurate modelling of canopy traits from seasonal Sentinel-2 imagery based on the vertical distribution of leaf traits. *ISPRS journal of photogrammetry and remote sensing,* 157, pp. 108-123.
Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S. and Gaudel-Vacaresse, A. (2017) Copernicus Sentinel-2 calibration and products validation status. *Remote Sensing,* 9 (6), pp. 584.
Ghimire, B., Rogan, J. and Miller, J. (2010) Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sensing Letters,* 1 (1), pp. 45-54.
Girardin, C.A.J., Malhi, Y., Aragao, L., Mamani, M., Huaraca Huasco, W., Durand, L., Feeley, K., Rapp, J., Silva-Espejo, J. and Silman, M. (2010) Net primary productivity allocation and cycling of carbon along a tropical forest elevational transect in the Peruvian Andes. *Global Change Biology,* 16 (12), pp. 3176-3192.
Gitelson, A.A., Gritz, Y. and Merzlyak, M.N. (2003) Relationships between leaf chlorophyll content and spectral reflectance and avaris
algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology,* 160 (3), pp. 271-282.
Gitelson, A.A. and Merzlyak, M.N. (1996) Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. *Journal of Plant Physiology,* 148 (3-4), pp. 494-500.
Graham, A.W. (2006) *The CSIRO rainforest permanent plots of North Queensland: site, structural, floristic and edaphic descriptions.* Atherton, CSIRO and Cooperative Research Centre for Tropical Rainforest.
Groenendijk, M., Dolman, A.J., Van der Molen, M., Leuning, R., Arneth, A., Delpierre, N., Gash, J., Lindroth, A., Richardson, A.D. and Verbeeck, H. (2011) Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data. *Agricultural and Forest Meteorology,* 151 (1), pp. 22-38.
Guanter, L., Alonso, L., Gómez-Chova, L., Amorós-López, J., Vila, J. and Moreno, J. (2007) Estimation of solar-induced vegetation fluorescence from space measurements. *Geophysical Research Letters,* 34 (8).
Gvozdevaite, A., Oliveras, I., Domingues, T.F., Peprah, T., Boakye, M., Afriyie, L., da Silva Peixoto, K., de Farias, J., Almeida de Oliveira, E. and Almeida Farias, C.C. (2018) Leaf-level photosynthetic capacity dynamics in relation to soil and foliar nutrients along forest–savanna boundaries in Ghana and Brazil. *Tree physiology,* 38 (12), pp. 1912-1925.
Halladay, K., Malhi, Y. and New, M. (2012) Cloud frequency climatology at the Andes/Amazon transition: 1. Seasonal and diurnal cycles. *Journal of Geophysical Research: Atmospheres,* 117.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J. and Loveland, T.R. (2013) High-resolution global maps of 21st-century forest cover change. *Science,* 342 (6160), pp. 850-853.
Hansen, M.C., Stehman, S.V. and Potapov, P.V. (2010) Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences,* 107 (19), pp. 8650-8655.
Haralick, R.M., Shanmugam, K. and Dinstein, I.H. (1973) Textural features for image classification. *IEEE transactions on systems, man, and cybernetics,* (6), pp. 610-621.
Haxeltine, A. and Prentice, I.C. (1996) A general model for the light-use efficiency of primary production. *Functional Ecology,* , pp. 551-561.
Heckmann, D., Schlüter, U. and Weber, A.P.M. (2017) Machine learning techniques for predicting crop photosynthetic capacity from leaf reflectance spectra. *Molecular plant,* 10 (6), pp. 878-890.
Hédl, R., Svátek, M., Dančák, M., Rodzay, A.W., Salleh AB, M. and Kamariah, A.S. (2009) A new technique for inventory of permanent plots in tropical forests: a case study from lowland dipterocarp forest in Kuala Belalong, Brunei Darussalam. *Blumea-Biodiversity, Evolution and Biogeography of Plants,* 54 (1-2), pp. 124-130.
Herrera, R., Jordan, C.F., Klinge, H. and Medina, E. (1978) Amazon ecosystems. Their structure and functioning with particular emphasis on nutrients. *Interciencia,* 3 (4), pp. 223-231.
Hölscher, D., Leuschner, C., Bohman, K., Hagemeier, M., Juhrbandt, J. and Tjitrosemito, S. (2006) Leaf gas exchange of trees in old-growth and young secondary forest stands in Sulawesi, Indonesia. *Trees,* 20 (3), pp. 278-285.
Houborg, R., Cescatti, A., Migliavacca, M. and Kustas, W.P. (2013) Satellite retrievals of leaf chlorophyll and photosynthetic capacity for improved modeling of GPP. *Agricultural and Forest Meteorology,* 177, pp. 10-23.
Huston, M.A. (2012) Precipitation, soils, NPP, and biodiversity: resurrection of Albrecht's curve. *Ecological Monographs,* 82 (3), pp. 277-296.
Ide, R., Nakaji, T. and Oguma, H. (2010) Assessment of canopy photosynthetic capacity and estimation of GPP by using spectral vegetation indices and the light–response function in a larch forest. *Agricultural and Forest Meteorology,* 150 (3), pp. 389-398.
Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F.W., Asner, G.P., Guralnick, R., Kattge, J., Latimer, A.M. and Moorcroft, P. (2016) Monitoring plant functional diversity from space. *Nature plants,* 2 (3), pp. 1-5.
Jin, P., Wang, Q., Iio, A. and Tenhunen, J. (2012) Retrieval of seasonal variation in photosynthetic capacity from multi-source vegetation indices. *Ecological Informatics,* 7 (1), pp. 7-18.
Joiner, J., Yoshida, Y., Vasilkov, A.P. and Middleton, E.M. (2011) First observations of global and seasonal terrestrial chlorophyll fluorescence from space. *Biogeosciences,* 8 (3), pp. 637-651.
Kattge, J. and Knorr, W. (2007) Temperature acclimation in a biochemical model of photosynthesis: a reanalysis of data from 36 species. *Plant, Cell & Environment,* 30 (9), pp. 1176-1190.
Kattge, J., Knorr, W., Raddatz, T. and Wirth, C. (2009) Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology,* 15 (4), pp. 976-991.
Kitajima, K., Mulkey, S.S. and Wright, S.J. (1997) Decline of photosynthetic capacity with leaf age in relation to leaf longevities for five tropical canopy tree species. *American Journal of Botany,* 84 (5), pp. 702-708.

Kitajima, K., Mulkey, S.S. and Wright, S.J. (1997) Seasonal leaf phenotypes in the canopy of a tropical dry forest: photosynthetic characteristics and associated traits. *Oecologia,* 109 (4), pp. 490-498.
Knorr, W. and Heimann, M. (2001) Uncertainties in global terrestrial biosphere modeling: 1. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. *Global Biogeochemical Cycles,* 15 (1), pp. 207-225.
Kokaly, R.F., Asner, G.P., Ollinger, S.V., Martin, M.E. and Wessman, C.A. (2009) Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment,* 113, pp. S78-S91.
Kotsiantis, S. and Pintelas, P. (2004) Combining bagging and boosting. *International Journal of Computational Intelligence,* 1 (4), pp. 324-333.
Kriegler, F.J., Malila, W.A., Nalepka, R.F. and Richardson, W. (1969) Preprocessing transformations and their effects on multispectral recognition. *Remote sensing of environment, VI,* , pp. 97.
Laurance, W.F. (2013) Emerging threats to tropical forests. In *Treetops at risk.* Springer, pp. 71-79.
Laurin, G.V., Puletti, N., Hawthorne, W., Liesenberg, V., Corona, P., Papale, D., Chen, Q. and Valentini, R. (2016) Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data. *Remote Sensing of Environment,* 176, pp. 163-176.
Lee, C.M., Cable, M.L., Hook, S.J., Green, R.O., Ustin, S.L., Mandl, D.J. and Middleton, E.M. (2015) An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. *Remote Sensing of Environment,* 167, pp. 6-19.
Lee, J.S.H., Jaafar, Z., Tan, A.K.J., Carrasco, L.R., Ewing, J.J., Bickford, D.P., Webb, E.L. and Koh, L.P. (2016) Toward clearer skies: Challenges in regulating transboundary haze in Southeast Asia. *Environmental Science & Policy,* 55, pp. 87-95.
Lloyd, J. and Farquhar, G.D. (2008) Effects of rising temperatures and [CO2] on the physiology of tropical forest trees. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 363 (1498), pp. 1811-1817.
Ma, X., Mahecha, M.D., Migliavacca, M., van der Plas, F., Benavides, R., Ratcliffe, S., Kattge, J., Richter, R., Musavi, T. and Baeten, L. (2019) Inferring plant functional diversity from space: the potential of Sentinel-2. *Remote Sensing of Environment,* 233, pp. 111368.
Malhi, Y., Gardner, T.A., Goldsmith, G.R., Silman, M.R. and Zelazowski, P. (2014) Tropical forests in the Anthropocene. *Annual Review of Environment and Resources,* 39, pp. 125-159.
Malhi, Y., Franklin, J., Seddon, N., Solan, M., Turner, M.G., Field, C.B. and Knowlton, N. (2020) *Climate change and ecosystems: Threats, opportunities and solutions.* The Royal Society.
Malhi, Y., Girardin, C., Metcalfe, D.B., Doughty, C.E., Aragão, L.E., Rifai, S.W., Oliveras, I., Shenkin, A., Aguirre-Gutiérrez, J. and Dahlsjö, C.A. (2021) The Global Ecosystems Monitoring network: Monitoring ecosystem productivity and carbon cycling across the tropics. *Biological Conservation,* 253, pp. 108889.
Medlyn, B.E., Badeck, F., De Pury, D., Barton, C., Broadmeadow, M., Ceulemans, R., De Angelis, P., Forstreuter, M., Jach, M.E. and Kellomäki, S. (1999) Effects of elevated [CO2] on photosynthesis in European forest species: a meta-analysis of model parameters. *Plant, Cell & Environment,* 22 (12), pp. 1475-1495.

Medvigy, D., Jeong, S., Clark, K.L., Skowronski, N.S. and Schäfer, K.,V.R. (2013) Effects of seasonal variation of photosynthetic capacity on the carbon fluxes of a temperate deciduous forest. *Journal of Geophysical Research: Biogeosciences,* 118 (4), pp. 1703-1714.
Meir, P., Grace, J. and Miranda, A.C. (2001) Leaf respiration in two tropical rainforests: constraints on physiology by phosphorus, nitrogen and temperature. *Functional Ecology,* 15 (3), pp. 378-387.
Meir, P., Levy, P.E., Grace, J. and Jarvis, P.G. (2007) Photosynthetic parameters from two contrasting woody vegetation types in West Africa. *Plant Ecology,* 192 (2), pp. 277-287.
Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R. and Moreno, J. (2009) Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. *Remote Sensing of Environment,* 113 (10), pp. 2037-2051.
Mitchard, E.T.A. (2018) The tropical forest carbon cycle and climate change. *Nature,* 559 (7715), pp. 527-534.
Muraoka, H., Noda, H.M., Nagai, S., Motohka, T., Saitoh, T.M., Nasahara, K.N. and Saigusa, N. (2013) Spectral vegetation indices as the indicator of canopy photosynthetic productivity in a deciduous broadleaf forest. *Journal of Plant Ecology,* 6 (5), pp. 393-407.
Nobre, C.A., Sellers, P.J. and Shukla, J. (1991) Amazonian deforestation and regional climate change. *Journal of Climate,* 4 (10), pp. 957-988.
Norby, R.J., Wullschleger, S.D., Gunderson, C.A., Johnson, D.W. and Ceulemans, R. (1999) Tree responses to rising CO2 in field experiments: implications for the future forest. *Plant, Cell & Environment,* 22 (6), pp. 683-714.
Oliveras, I., Bentley, L., Fyllas, N.M., Gvozdevaite, A., Shenkin, A.F., Peprah, T., Morandi, P., Peixoto, K.S., Boakye, M. and Adu-Bredu, S. (2020) The influence of taxonomy and environment on leaf trait variation along tropical abiotic gradients. *Frontiers in Forests and Global Change,* 3, pp. 18.
Osuna, J.L., Baldocchi, D.D., Kobayashi, H. and Dawson, T.E. (2015) Seasonal trends in photosynthesis and electron transport during the Mediterranean summer drought in leaves of deciduous oaks. *Tree physiology,* 35 (5), pp. 485-500.
Philippon, N., Cornu, G., Monteil, L., Gond, V., Moron, V., Pergaud, J., Sèze, G., Bigot, S., Camberlin, P. and Doumenge, C. (2019) The light-deficient climates of western Central African evergreen forests. *Environmental Research Letters,* 14 (3), pp. 034007.
Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G. and Bayol, N. (2020) Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature communications,* 11 (1), pp. 1-11.
Prentice, I.C., Dong, N., Gleason, S.M., Maire, V. and Wright, I.J. (2014) Balancing the costs of carbon gain and water transport: testing a new theoretical framework for plant functional ecology. *Ecology Letters,* 17 (1), pp. 82-91.
Propastin, P., Ibrom, A., Knohl, A. and Erasmi, S. (2012) Effects of canopy photosynthesis saturation on the estimation of gross primary productivity from MODIS data in a tropical forest. *Remote Sensing of Environment,* 121, pp. 252-260.
Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H. and Sorooshian, S. (1994) A modified soil adjusted vegetation index. *Remote Sensing of Environment,* 48 (2), pp. 119-126.
Quesada, C.A., Lloyd, J., Schwarz, M., Patiño, S., Baker, T.R., Czimczik, C., Fyllas, N.M., Martinelli, L., Nardoto, G.B. and Schmerler, J. (2010) Variations in chemical and physical properties of Amazon forest soils in relation to their genesis. *Biogeosciences,* 7 (5), pp. 1515-1541.

Quesada, C.A., Hodnett, M.G., Breyer, L.M., Santos, A.J.B., Andrade, S., Miranda, H.S., Miranda, A.C. and Lloyd, J. (2008) Seasonal variations in soil water in two woodland savannas of central Brazil with different fire histories. *Tree physiology,* 28 (3), pp. 405-415.

Raaimakers, D., Boot, R., Dijkstra, P. and Pot, S. (1995) Photosynthetic rates in relation to leaf phosphorus content in pioneer versus climax tropical rainforest trees. *Oecologia,* 102 (1), pp. 120-125.

Reich, P.B., Walters, M.B., Tjoelker, M.G., Vanderklein, D. and Buschena, C. (1998) Photosynthesis and respiration rates depend on leaf and root morphology and nitrogen concentration in nine boreal tree species differing in relative growth rate. *Functional Ecology,* 12 (3), pp. 395-405.

Rogers, A. (2014) The use and misuse of V c, max in Earth System Models. *Photosynthesis Research,* 119 (1), pp. 15-29.

Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M., Reeves, M. and Hashimoto, H. (2004) A continuous satellite-derived measure of global terrestrial primary production. *Bioscience,* 54 (6), pp. 547-560.

Salazar, L.F., Nobre, C.A. and Oyama, M.D. (2007) Climate change consequences on the biome distribution in tropical South America. *Geophysical Research Letters,* 34 (9).

Sanchez, A.H., Picoli, M.C.A., Camara, G., Andrade, P.R., Chaves, M.E.D., Lechler, S., Soares, A.R., Marujo, R.F.B., Simões, R.E.,O. and Ferreira, K.R. (2020) Comparison of Cloud cover detection algorithms on sentinel–2 images of the amazon tropical forest. *Remote Sensing,* 12 (8), pp. 1284.

Schmitt, S., Hérault, B., Ducouret, É., Baranger, A., Tysklind, N., Heuertz, M., Marcon, É., Cazal, S.O. and Derroire, G. (2020) Topography consistently drives intra-and inter-specific leaf trait variation within tree species complexes in a Neotropical forest. *Oikos,* 129 (10), pp. 1521-1530.

Serbin, S.P., Dillaway, D.N., Kruger, E.L. and Townsend, P.A. (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *Journal of experimental botany,* 63 (1), pp. 489-502.

Serbin, S.P., Singh, A., Desai, A.R., Dubois, S.G., Jablonski, A.D., Kingdon, C.C., Kruger, E.L. and Townsend, P.A. (2015) Remotely estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using imaging spectroscopy. *Remote Sensing of Environment,* 167, pp. 78-87.

Serbin, S.P., Singh, A., McNeil, B.E., Kingdon, C.C. and Townsend, P.A. (2014) Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecological Applications,* 24 (7), pp. 1651-1669.

Shenkin, A., Bentley, L.P., Oliveras, I., Salinas, N., Adu-Bredu, S., Marimon-Junior, B., Marimon, B.S., Peprah, T., Choque, E.L. and Trujillo Rodriguez, L. (2020) The influence of ecosystem and phylogeny on tropical tree crown size and shape. *Frontiers in Forests and Global Change,* 3, pp. 109.

Smith, N.G., Keenan, T.F., Colin Prentice, I., Wang, H., Wright, I.J., Niinemets, Ü., Crous, K.Y., Domingues, T.F., Guerrieri, R. and Yoko Ishida, F. (2019) Global photosynthetic capacity is optimized to the environment. *Ecology Letters,* 22 (3), pp. 506-517.

Swinfield, T., Both, S., Riutta, T., Bongalov, B., Elias, D., Majalap-Lee, N., Ostle, N., Svátek, M., Kvasnica, J. and Milodowski, D. (2020) Imaging spectroscopy reveals the effects of topography and logging on the leaf chemistry of tropical forest canopy trees. *Global Change Biology,* 26 (2), pp. 989-1002.

Thomson, E.R., Malhi, Y., Bartholomeus, H., Oliveras, I., Gvozdevaite, A., Peprah, T., Suomalainen, J., Quansah, J., Seidu, J. and Adonteng, C. (2018) Mapping the leaf economic spectrum across West African tropical forests using UAV-Acquired hyperspectral imagery. *Remote Sensing,* 10 (10), pp. 1532.
Thomson, E.R., Spiegel, M.P., Althuizen, I.H.J., Bass, P., Chen, S., Chmurzynski, A., Halbritter, A.H., Henn, J.J., Jónsdóttir, I.,S. and Klanderud, K. (2021) Multiscale mapping of plant functional groups and plant traits in the High Arctic using field spectroscopy, UAV imagery and Sentinel-2 data. *Environmental Research Letters,* 16 (5), pp. 055006.
Turner, D.P., Ritts, W.D., Cohen, W.B., Gower, S.T., Running, S.W., Zhao, M., Costa, M.H., Kirschbaum, A.A., Ham, J.M. and Saleska, S.R. (2006) Evaluation of MODIS NPP and GPP products across multiple biomes. *Remote Sensing of Environment,* 102 (3-4), pp. 282-292.
Ustin, S.L., Gitelson, A.A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J.A. and Zarco-Tejada, P. (2009) Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sensing of Environment,* 113, pp. S67-S77.
Valladares, F. and Niinemets, Ü. (2008) Shade tolerance, a key plant feature of complex nature and consequences. *Annual Review of Ecology, Evolution, and Systematics,* 39, pp. 237-257.
Vårhammar, A., Wallin, G., McLean, C.M., Dusenge, M.E., Medlyn, B.E., Hasper, T.B., Nsabimana, D. and Uddling, J. (2015) Photosynthetic temperature responses of tree species in Rwanda: evidence of pronounced negative effects of high temperature in montane rainforest climax species. *New Phytologist,* 206 (3), pp. 1000-1012.
Wadoux, A.M.J.-., Heuvelink, G.B.M., De Bruin, S. and Brus, D.J. (2021) Spatial cross-validation is not the right way to evaluate map accuracy. *Ecological Modelling,* 457, pp. 109692.
Walker, A.P., Beckerman, A.P., Gu, L., Kattge, J., Cernusak, L.A., Domingues, T.F., Scales, J.C., Wohlfahrt, G., Wullschleger, S.D. and Woodward, F.I. (2014) The relationship of leaf photosynthetic traits–Vcmax and Jmax–to leaf nitrogen, leaf phosphorus, and specific leaf area: a meta-analysis and modeling study. *Ecology and evolution,* 4 (16), pp. 3218-3235.
Wright, I.J., Reich, P.B. and Westoby, M. (2003) Least-cost input mixtures of water and nitrogen for photosynthesis. *The American Naturalist,* 161 (1), pp. 98-111.
Xu, L. and Baldocchi, D.D. (2003) Seasonal trends in photosynthetic parameters and stomatal conductance of blue oak (Quercus douglasii) under prolonged summer drought and high temperature. *Tree physiology,* 23 (13), pp. 865-877.
Yanhong, T., Naoki, K., Akio, F. and Awang, M. (1996) Light reduction by regional haze and its effect on simulated leaf photosynthesis in a tropical forest of Malaysia. *Forest Ecology and Management,* 89 (1-3), pp. 205-211.
Zaehle, S., Sitch, S., Smith, B. and Hatterman, F. (2005) Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics. *Global Biogeochemical Cycles,* 19 (3).

8.10. Supplementary material

An introduction to the GEM network

The Global Ecosystems Monitoring (GEM) network is a network of sites that uses a standard protocol to track the productivity and carbon cycling of terrestrial ecosystems (Malhi *et al*. 2021). From its origins focused on Amazonian and Andean forests, since 2010, GEM has expanded to cover all tropical continents, and also includes a range of tropical savanna and temperate forest sites. To date, the network spans 294 plots covering 178 ha (**Figure S8.1**).

This chapter included data from plots located in Australia, Brazil, Gabon, Ghana, Maysian Borneo and Peru (**Table S8.1**).

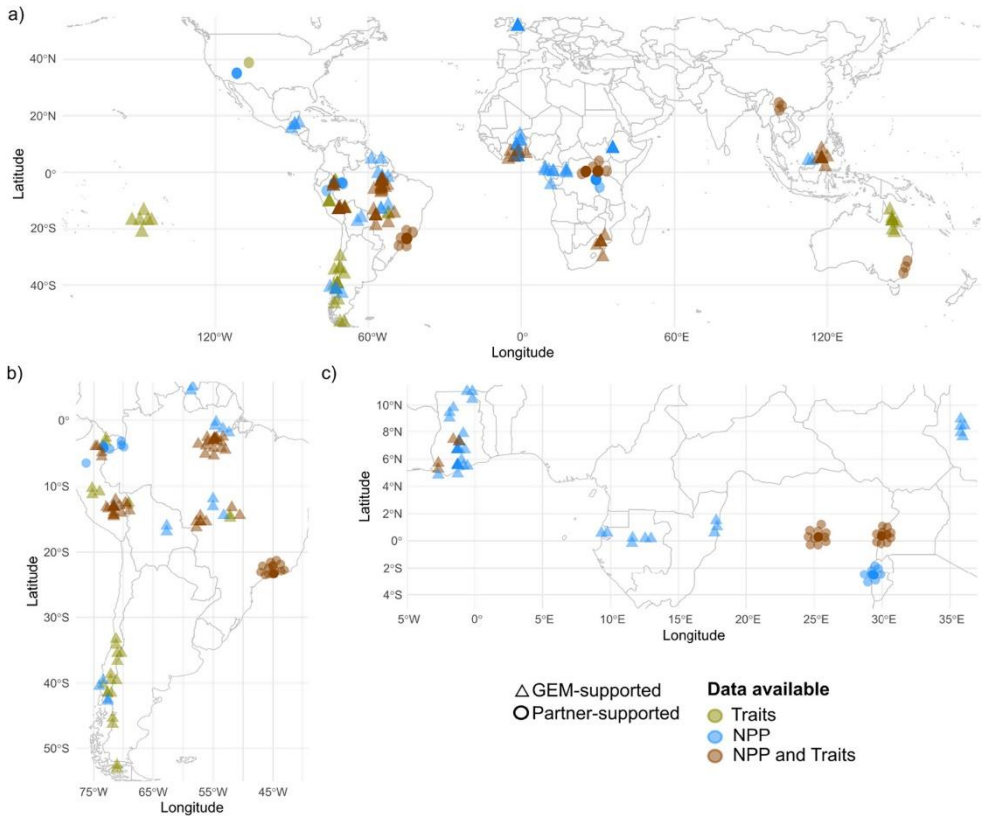


Figure S8.1: The GEM Network, as of early 2020. Open circles indicate sites where carbon cycle data alone have been collected; grey filled circles indicate where plant traits data and forest inventory data alone have been collected; filled circles where both detailed carbon cycle and plant functional traits data have been collected. Circles indicate sites that are centrally supported by the GEM network; triangles indicate sites primarily supported by external partners. Reproduced from Malhi *et al*. (2021).

Australia

Six GEM plots are located in Queensland Australia, around Cairns and Cape Tribulation (Graham *et al.* 2006). They span an elevation gradient from 15-1140 m and are located in a mixture of moist tropical forest and tropical montane forest. Climatically, the plots span wet lowlands (AEP-41; DRO-01), dry and seasonal uplands (AEP-02; AEP-33) and wet highlands (ROB-06; AEP-03). Soils are generally deeply weathered granite, with the exception of AEP-33, which is situated on nutrient-rich, unweathered basalt soils. The disturbance history of the plots is varied, with evidence of selective logging present and/or natural disturbance in nearly all plots, except AEP-03.

Brazil, Nova Xavantina

Three GEM plots are located in, or near, the Bacaba Municipal Park, Nova Xavantina, eastern Mato Grosso State, Central Brazil. The park is located in the Cerrado-Amazon transition region and spans a tall closed canopy forest plot (VCR-02), a transitional (cerradão) forest plot (NXV-02) and a savanna forest plot (NXV-01). The climate is characterized by two well-defined seasons: hot and wet from October to March, and cool and dry from April to September. Annual rainfall is 1440 mm and the average temperature is 24.8 °C. Altitude is approximately 300 m. The soils of plots are plinthosols and ferralsols (Gvozdevaite *et al*. 2018).

Brazil, Santarem

Three GEM plots are situated near Santarem, western Pará State, Brazil. Average temperatures vary between 25 °C – 28 °C, with relative humidity at 86%. Annual precipitation is around 1900 mm. The region has a marked dry season from August to November, when precipitation is less than 100 mm per month. The plots span a disturbance gradient, from moist evergreen old-growth forest (ST0-03) to logged forest (STQ-08) and previously logged and burned forest (STQ-11). Soils are clay-rich Oxisols (see Berenguer *et al*. 2018 for more details).

Gabon

Three GEM plots are located in Gabon. One plot is situated in coastal forest in Mondah Forest Reserve, north of Libreville (MNG-04), and the other two plots are situated inland, in Lopé National Park, home to forest elephants and other megafauna (LPG-01; LPG-2). Mean annual temperature across the plots is around 26 °C. Annual precipitation is much higher on the coast (> 3000 mm) and lower inland (~1700 mm). Soils on the coast are tertiary sands, while the soils of Lopé are granite-derived Oxisols.

Peru

Nine GEM plots in Peru form an Amazon-Andes elevation gradient, stretching from 200-3500 m.a.s.l. (Malhi *et al*. 2017). Two plots are located in the Amazonian lowlands (<300 m: TAM-05; TAM-06); two plots are located on the Pantiacolla front range of the Andes (600–900 m: PAN-02; PAN-03); and five plots are located up the eastern slope of the Andes, in the Kosñipata valley (1,500–3,500 m: SPD-01; SPD-02; ESP-01; WAY-01; ACJ-01). Mean annual air temperatures over the gradient ranges from 25.28 ºC (223 m) to 10.68 ºC (3536 m). There is a close association between elevation and temperature. Precipitation shows a humped relationship with elevation, with highest recorded values (3309 mm) near the base of the mountain range (1000 - 1500 m). The cloud immersion zone begins around 1500 - 1800 m and extends to just before the tree line, at approximately 3200-3400 m. Above the tree line, the forest makes a sharp transition to fire-dominated puna grasslands. Plots below 1000 m are situated on clay rich sediments formed from alluvial deposition. The lowland Tambopata plots are located on nutrient poor 'white sand' soils. Above 2000 m, soils are highly acidic (pH 3.6) but soil carbon, nitrogen and phosphorus values are much higher, concentrated in a thick layer of humic material above carbon- rich topsoil (Girardin *et al*. 2010).

Plot Code	Country	Longitude	Latitude	Biome	Elevation (m)	Temp (°C)	MAT (mm)	Solar (W/m^2)	PET (mm)	CWD (mm)
AEP-02	Australia	145.5869	-17.1466	Montane tropical forest	710	20.88	1530	225.1	122	27.9
AEP-03	Australia	145.5936	-17.0875	Montane tropical forest	1140	19.93	1432	224.5	117.3	28.6
AEP-33	Australia	145.5711	-17.2848	Montane tropical forest	730	21.09	2320	222.8	122.7	16
AEP-41	Australia	145.4411	-16.1361	Tropical rainforest	15	24.07	3112	224.8	125.6	14.5
DRO-01	Australia	145.4303	-16.1033	Tropical rainforest	58	23.3	2957	221.6	121.8	12.3
ROB-06	Australia	145.6339	-17.1210	Montane tropical forest	720	20.19	2238	221.9	116.8	17.8
NXV-01	Brazil	-52.2103	-14.4227	Savannah	295	25.01	1580	186	108.8	30.4
NXV-02	Brazil	-52.2104	-14.4227	Savannah	295	25.01	1580	186	108.8	30.4
VCR-02	Brazil	-52.1680	-14.8320	Tropical dry forest	289	25.05	1589	186.9	109.4	31.1
STO-03	Brazil	-54.9633	-3.2963	Tropical rainforest	133	26.06	1840	171.4	98.1	14.9
STQ-08	Brazil	-54.6749	-2.5146	Tropical rainforest	135	26.53	2058	175.4	101	15.2
STQ-11	Brazil	-54.6708	-2.5736	Tropical rainforest	114	26.04	1997	175.1	99.4	14.6
LPG-01	Gabon	11.5737	-0.1741	Tropical rainforest	399	26.59	1701	142.7	87.8	7.6
LPG-02	Gabon	11.6152	-0.2157	Tropical rainforest	288	25.58	1746	143	85.1	7.2
MNG-04	Gabon	9.3237	0.5767	Tropical rainforest	39	26.06	3043	172.1	98	8.6
ACJ-01	Peru	-71.6323	-13.1469	Montane tropical forest	3537	10.68	567	161	83.4	39.3
ESP-01	Peru	-71.5948	-13.1751	Montane tropical forest	2863	12.13	455	159	84.6	48.7
PAN-02	Peru	-71.2627	-12.6496	Montane tropical forest	595	23.94	3309	160.3	93.8	3.8
PAN-03	Peru	-71.2745	-12.6383	Montane tropical forest	848	23.94	3309	160.3	93.8	3.8
SPD-01	Peru	-71.5423	-13.0475	Montane tropical forest	1776	18.74	1641	156.9	87.9	9.1
SPD-02	Peru	-71.5365	-13.0491	Montane tropical forest	1527	19.57	1920	157.5	89.2	7.6
TAM-05	Peru	-69.2705	-12.8309	Tropical rainforest	223	25.28	2436	167.6	98.3	5.2
TAM-06	Peru	-69.2960	-12.8385	Tropical rainforest	215	25.31	2496	167.2	98.2	5
WAY-01	Peru	-71.5874	-13.1908	Montane tropical forest	3045	12.13	455	159	84.6	48.7

Table S8.1: Metadata of sites in the GEM network that appear in this thesis. The column headings indicate Plot Code; Country; Latitude; Longitude; Biome; Elevation (m a.s.l.); Air Temperature (Celsius); Annual Precipitation (mm); Mean Solar Radiation (W/m^2); Annual Potential Evapotranspiration (mm); Maximum Cumulative Water Deficit (mm). Adapted from Malhi *et al.* (2021).

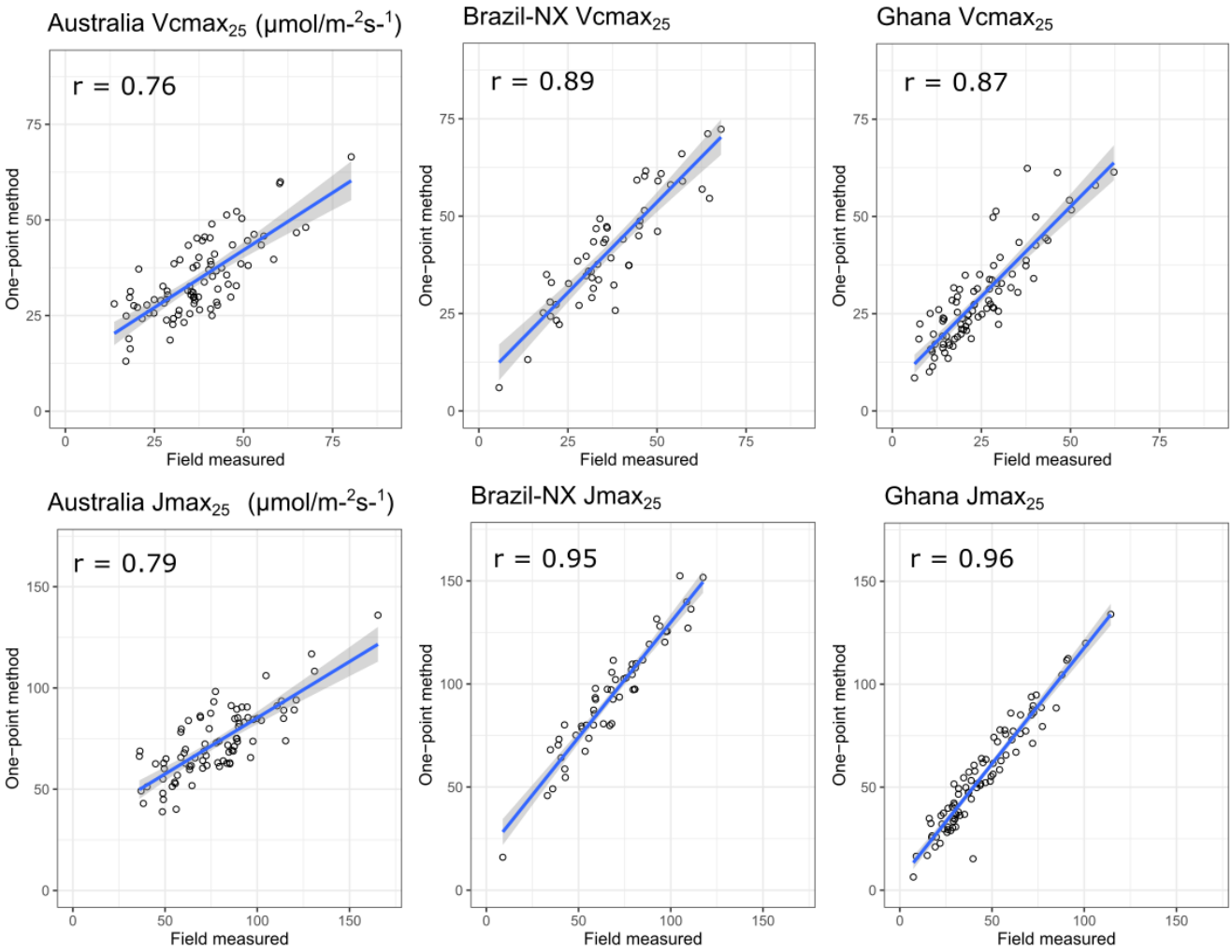


Figure S8.2*:* Prediction accuracy of de Kauwe *et al.*'s (2016) 'one-point' method compared against $Vcmax_{25}$ and $Jmax_{25}$ derived from field-measured A-Ci-curves.

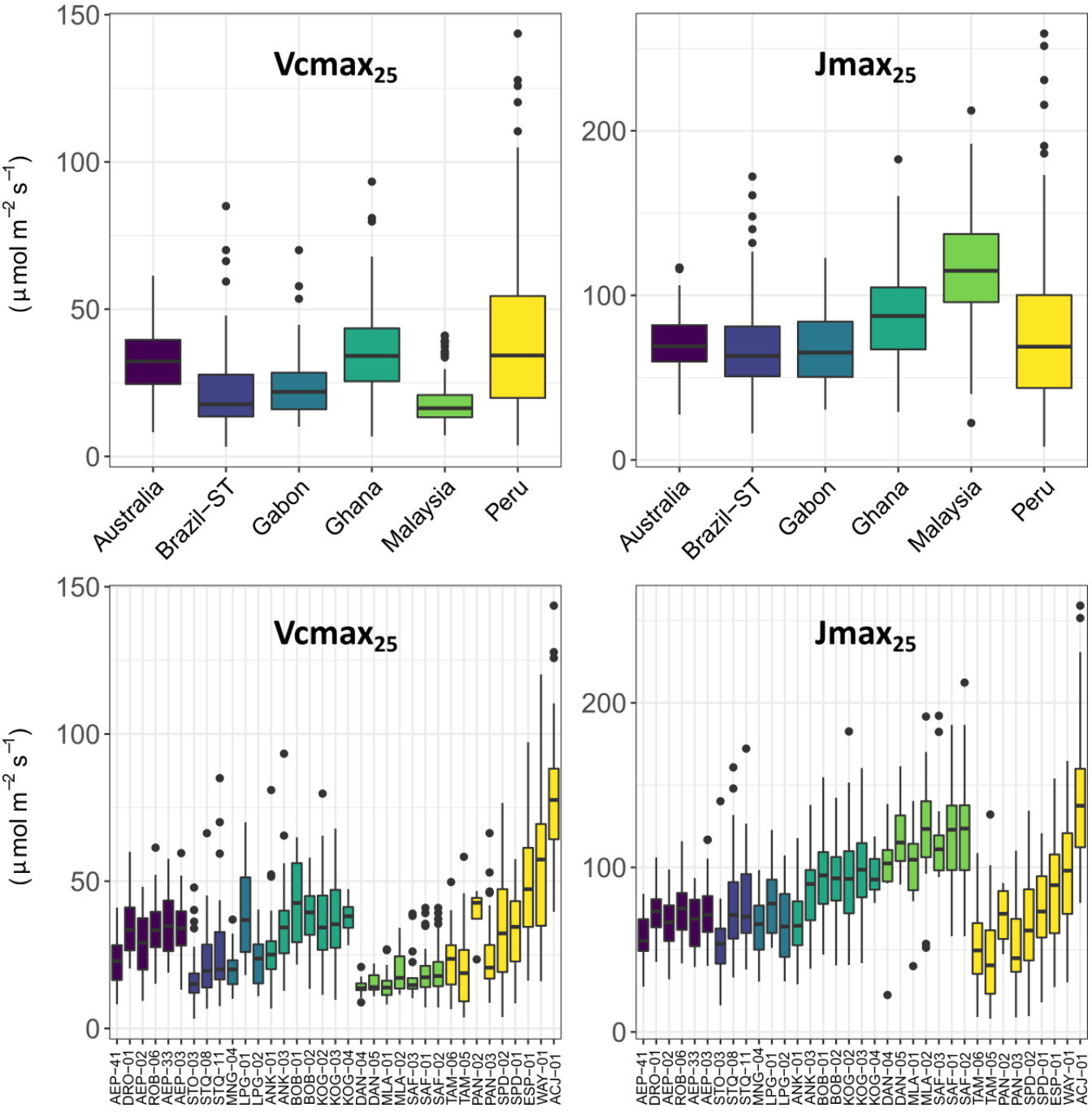


Figure S8.3: Field-sampled $Vcmax_{25}$ and $Jmax_{25}$ values, averaged per tree crown. $Vcmax_{25}$ and $Jmax_{25}$ values were derived using the one-point method (de Kauwe *et al.* 2016). More information on plot characteristics can be found in **Table S8.1**.

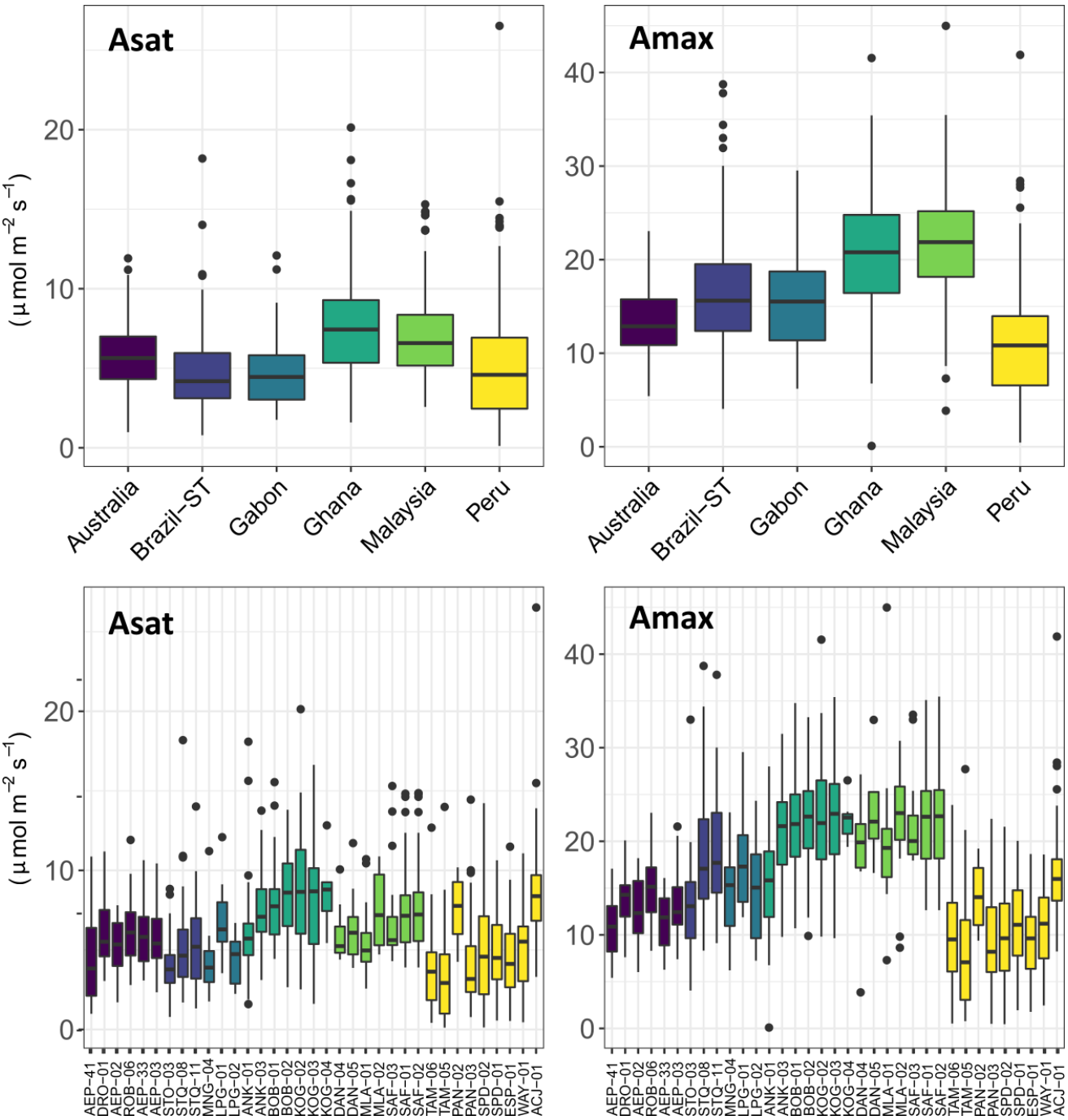


Figure S8.4: Field-sampled Asat and Amax values, averaged per tree crown. (*Top*) Comparison of Asat and Amax values across 6 pan-tropical sites. (*Bottom*) Comparison of Asat and Amax values across 35 tropical forest plots, coloured by site. More information on plot characteristics can be found in **Table S8.1**.

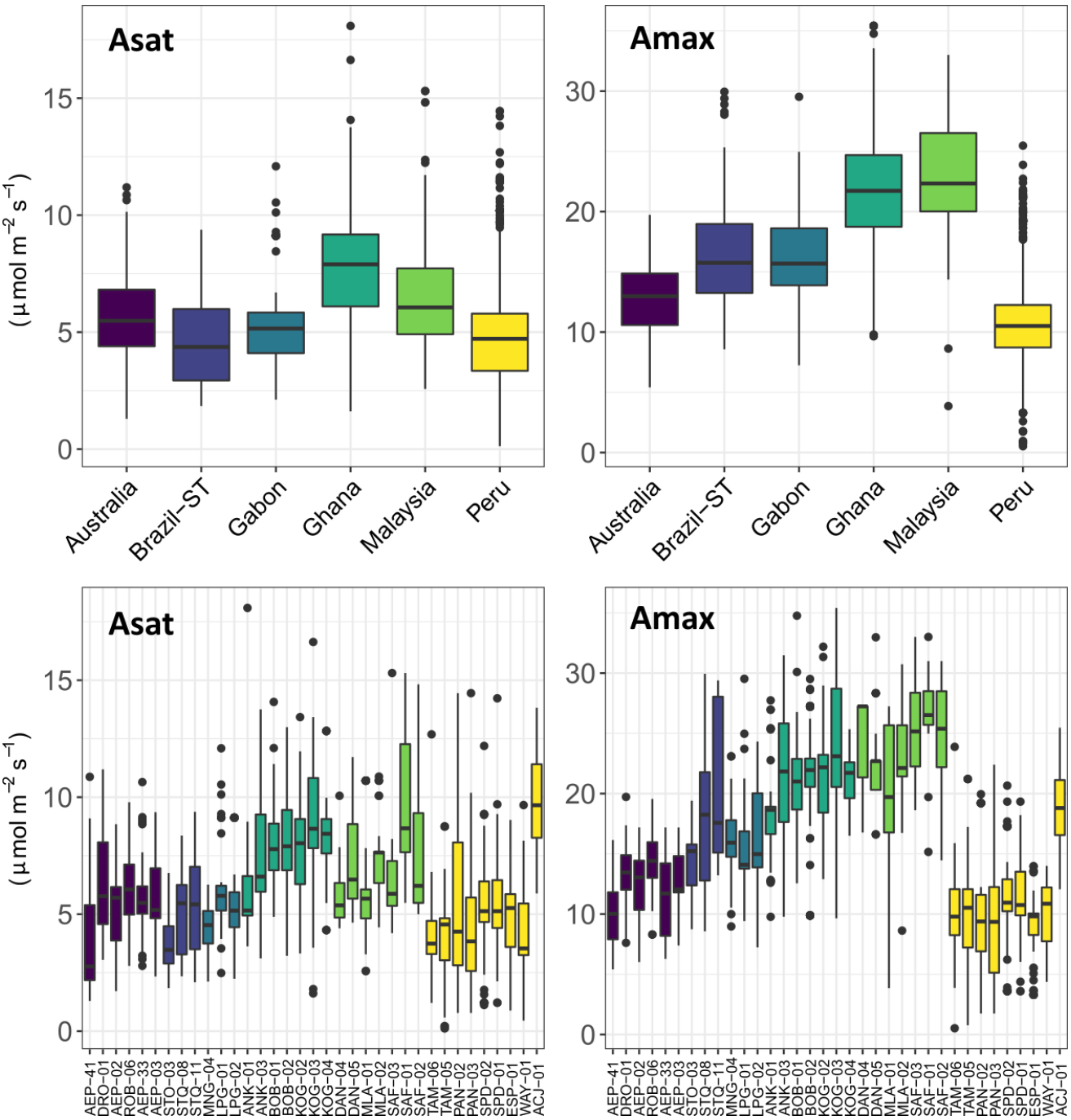


Figure S8.5: Trait values represent 10x10m pixel-level community-weighted mean of the most visible sun foliage. (*Top*) Comparison of Asat and Amax values across 6 pan-tropical sites. (*Bottom*) Comparison of Asat and Amax values across 35 tropical forest plots, coloured by site. More information on plot characteristics can be found in **Table S8.1**.

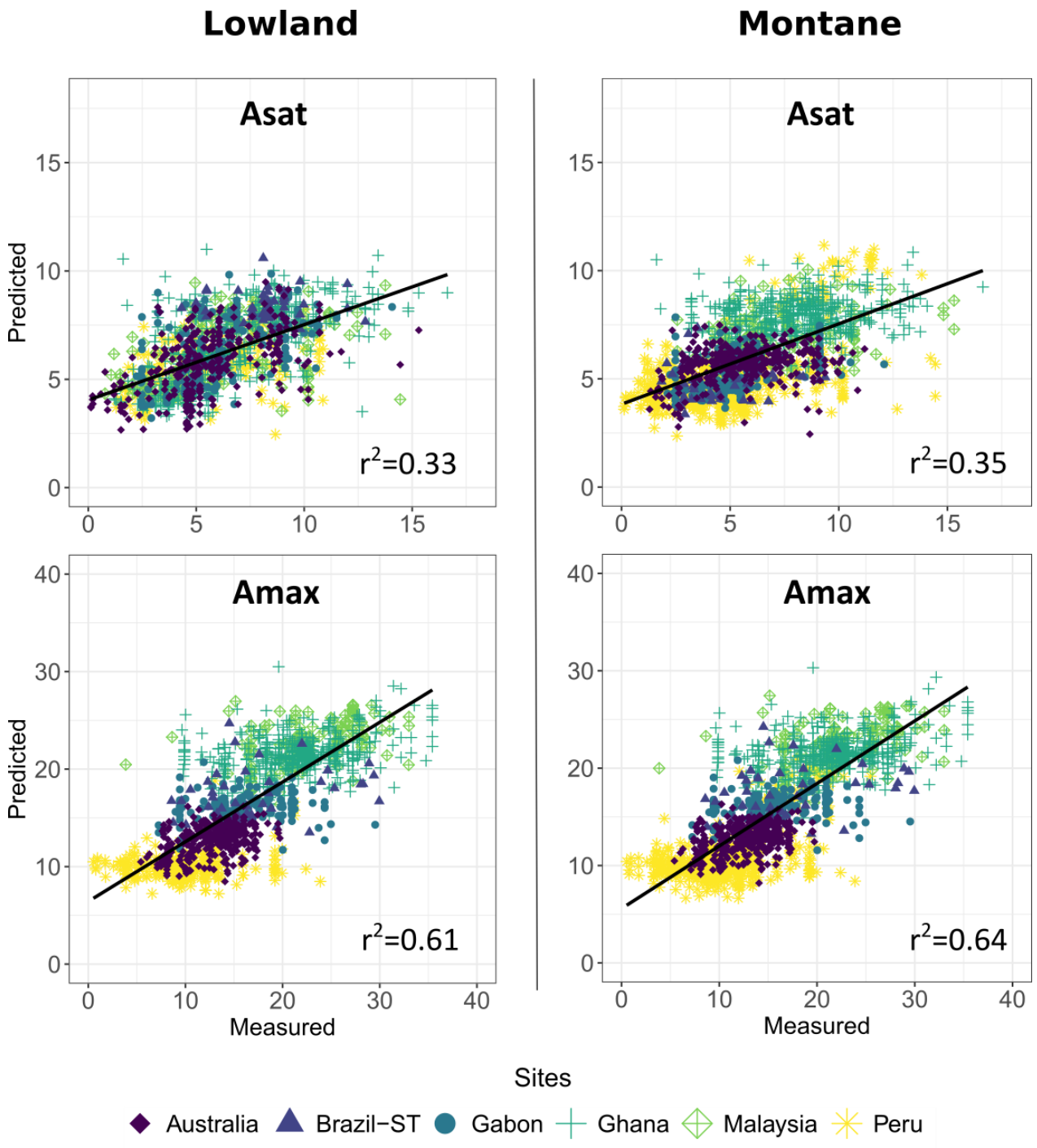


Figure S8.6: Pixel-level Asat and Amax values plotted against model predictions for lowland and montane forest. Points represent random forest regression results for the 30% 'out of bag' validation dataset. Points are coloured according to site. The black line shows the 1:1 relationship between observed and predicted values. Model prediction accuracy is shown in the bottom left. Full model results can be found in **Table S8.2.**

	Lowland Forest					
	Asat			Amax		
	R-Squared	*RMSE*	*%RMSE*	*R-Squared*	*RMSE*	*%RMSE*
All plots	0.33	2.14	35	0.61	4.03	24
Spatial folds	0.15	2.57	41	0.28	6.6	40
Australia	0.14	1.84	33	0.22	2.69	21
Brazil-ST	0.20	1.76	38	0.28	5.65	32
Gabon	0.25	1.55	30	0.09	3.60	23
Ghana	0.10	2.32	30	0.12	4.62	22
Malaysia	0.24	2.13	32	0.26	3.89	17
Peru	0.07	2.37	54	0.05	4.26	43
	Montane Forest					
All plots	0.35	2.09	35	0.64	3.87	25
Spatial folds	0.12	2.67	44	0.27	6.17	38
Australia	0.12	1.83	33	0.20	2.62	21
Peru	0.34	2.13	43	0.32	3.6	33

Table S8.2: Statistical results of predictive models for Asat and Amax created using random forest regression. All statics represent the average of 100 model runs. Only data from lowlands plots (>1000m) was used to train lowland forest models. All plot data were used to train montane forest models. Pan-tropical models were validated twice. (1) using a random 70:30 training/validation split where all pixels were treated as independent samples. (2) splitting the plots into spatial folds to ensure that no pixels from the same, or nearby, plots used in the training dataset entered the validation dataset. See **Table 8.1** for spatial fold groupings. Models were also run at the site level using the site-specific data.

Single Sentinel Images	Lowland Forest					
	$Vcmax_{25}$			$Jmax_{25}$		
	R-Squared	*RMSE*	*%RMSE*	*R-Squared*	*RMSE*	*%RMSE*
All plots	*0.36*	10.24	35	*0.57*	19.90	25
Spatial folds	*0.14*	12.50	44	*0.29*	29.95	39
Australia	*0.16*	8.67	28	*0.18*	13.28	20
Brazil-ST	*0.32*	8.85	42	*0.48*	22.44	31
Gabon	*0.30*	9.48	34	*0.09*	15.50	23
Ghana	*0.11*	11.13	31	*0.18*	21.10	23
Malaysia	*0.52*	4.62	27	*0.55*	17.30	14
Peru	*0.04*	13.50	56	*0.1*	26.16	49

Table S8.3: Statistical results of predictive models for $Vcmax_{25}$ and $Jmax_{25}$ created using random forest regression. The same as **Table 8.3**, but using single Sentinel images captured as close in time to photosynthesis sampling as possible, instead of a 2019 Sentinel composite. All single Sentinel images were manually checked to exclude cloud cover and cloud shadow.

Single Sentinel Images	Lowland Forest					
	Asat			Amax		
	R-Squared	*RMSE*	*%RMSE*	*R-Squared*	*RMSE*	*%RMSE*
All plots	*0.36*	2.25	37	*0.60*	4.33	26
Spatial folds	*0.2*	2.65	42	*0.21*	6.14	36
Australia	*0.17*	1.81	33	*0.25*	2.73	22
Brazil-ST	*0.20*	1.65	39	*0.44*	5.53	32
Gabon	*0.15*	1.63	33	*0.10*	3.56	23
Ghana	*0.13*	2.28	29	*0.17*	4.61	22
Malaysia	*0.48*	1.81	26	*0.47*	3.05	13
Peru	*0.03*	3.03	72	*0.11*	5.65	55

Table S8.4*:* Statistical results of predictive models for Asat and Amax created using random forest regression. The same as **Table S8.2**, but using single Sentinel images captured as close in time to photosynthesis sampling as possible, instead of a 2019 Sentinel composite. All single Sentinel images were manually checked to exclude cloud cover and cloud shadow.

Only spectral information	Lowland Forest					
	$Vcmax_{25}$			$Jmax_{25}$		
	R-Squared	*RMSE*	*%RMSE*	*R-Squared*	*RMSE*	*%RMSE*
All plots	*0.30*	9.88	34	*0.34*	23.40	30
Spatial folds	*0.12*	13.30	47	*0.09*	33.33	42
Australia	*0.15*	8.80	28	*0.15*	13.00	19
Brazil-ST	*0.21*	7.97	38	*0.29*	21.19	30
Gabon	*0.34*	10.25	35	*0.08*	16.43	23
Ghana	*0.05*	10.41	29	*0.10*	21.14	24
Malaysia	*0.23*	5.37	32	*0.19*	23.80	19
Peru	*0.09*	9.74	40	*0.06*	18.64	36

Table S8.5*:* Statistical results of predictive models for $Vcmax_{25}$ and $Jmax_{25}$ created using random forest regression. The same as **Table 8.3**, but using excluding environmental information and only using spectral, vegetation and textual predictors, extracted or computed from Sentinel imagery.

Only spectral information	**Lowland Forest**					
	Asat			**Amax**		
	R-Squared	*RMSE*	*%RMSE*	*R-Squared*	*RMSE*	*%RMSE*
All plots	*0.24*	2.24	37	*0.42*	4.95	30
Spatial folds	*0.17*	2.61	43	*0.21*	6.69	40
Australia	*0.13*	1.86	33	*0.17*	2.72	22
Brazil-ST	*0.28*	1.71	37	*0.32*	5.20	30
Gabon	*0.26*	1.54	29	*0.10*	3.59	23
Ghana	*0.08*	2.33	30	*0.09*	4.68	22
Malaysia	*0.14*	2.24	33	*0.19*	4.28	19
Peru	*0.07*	2.38	54	*0.05*	4.26	43

Table S8.6*:* Statistical results of predictive models for Asat and Amax created using random forest regression. The same as **Table 8.3**, but using excluding environmental information and only using spectral, vegetation and textual predictors, extracted or computed from Sentinel imagery.

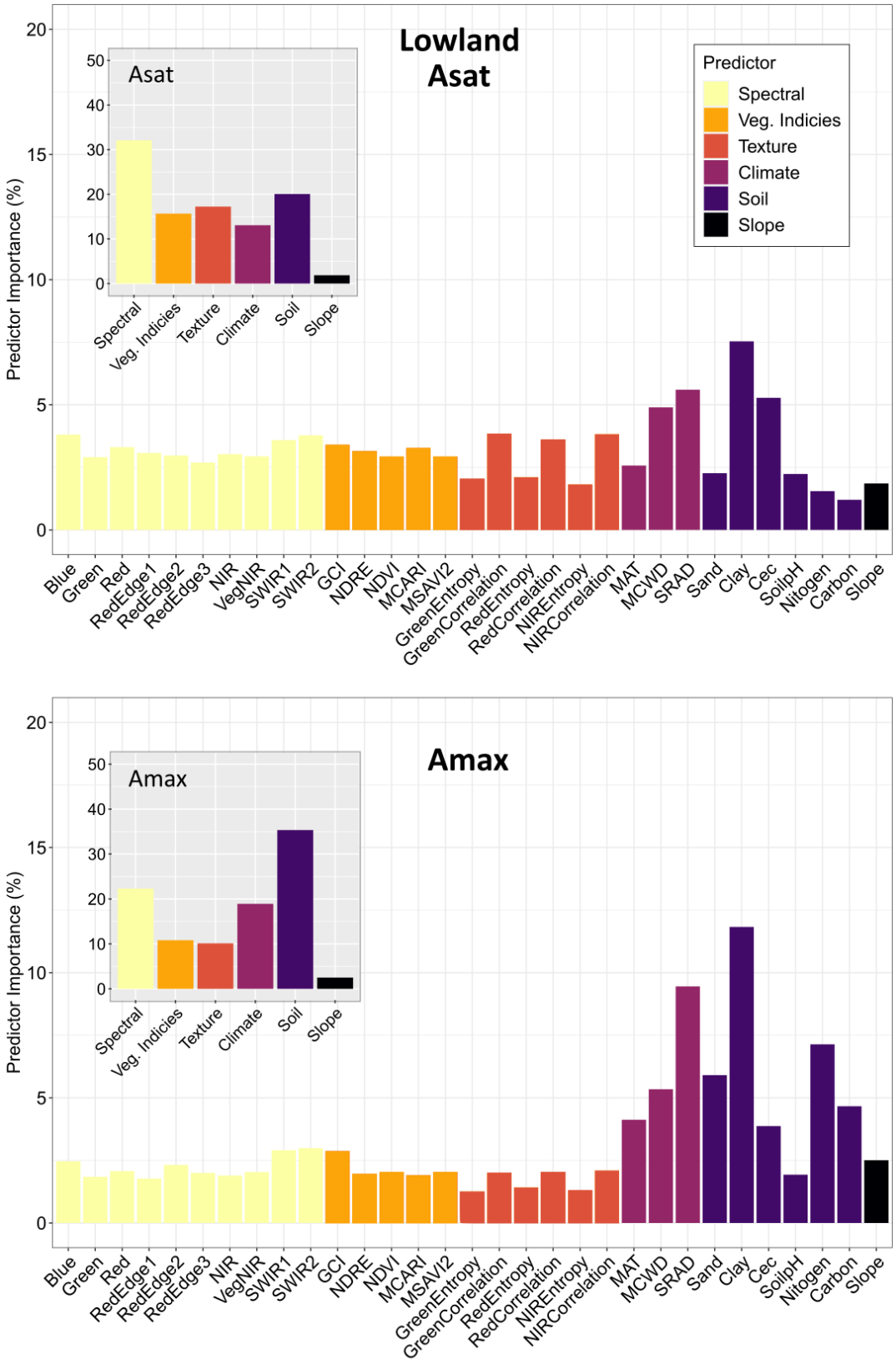


Figure S8.7: Percentage importance of each spectral, texture, climatic and soil variable for predicting lowland Asat and Amax. *Insert*: Total importance for each predictor group, represented as a percentage.

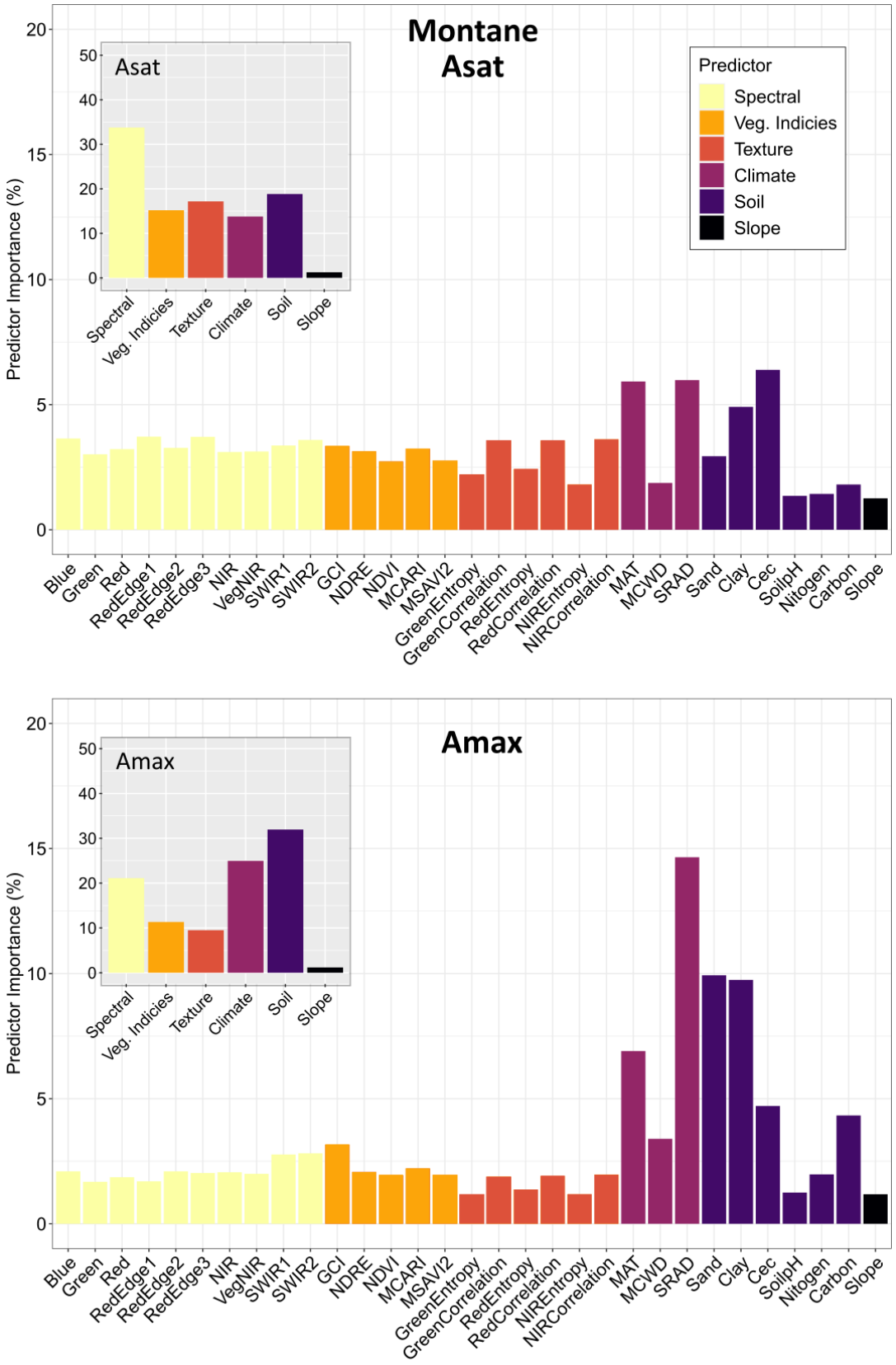


Figure S8.8: Percentage importance of each spectral, texture, climatic and soil variable for predicting montane Asat and Amax. *Insert*: Total importance for each predictor group, represented as a percentage.

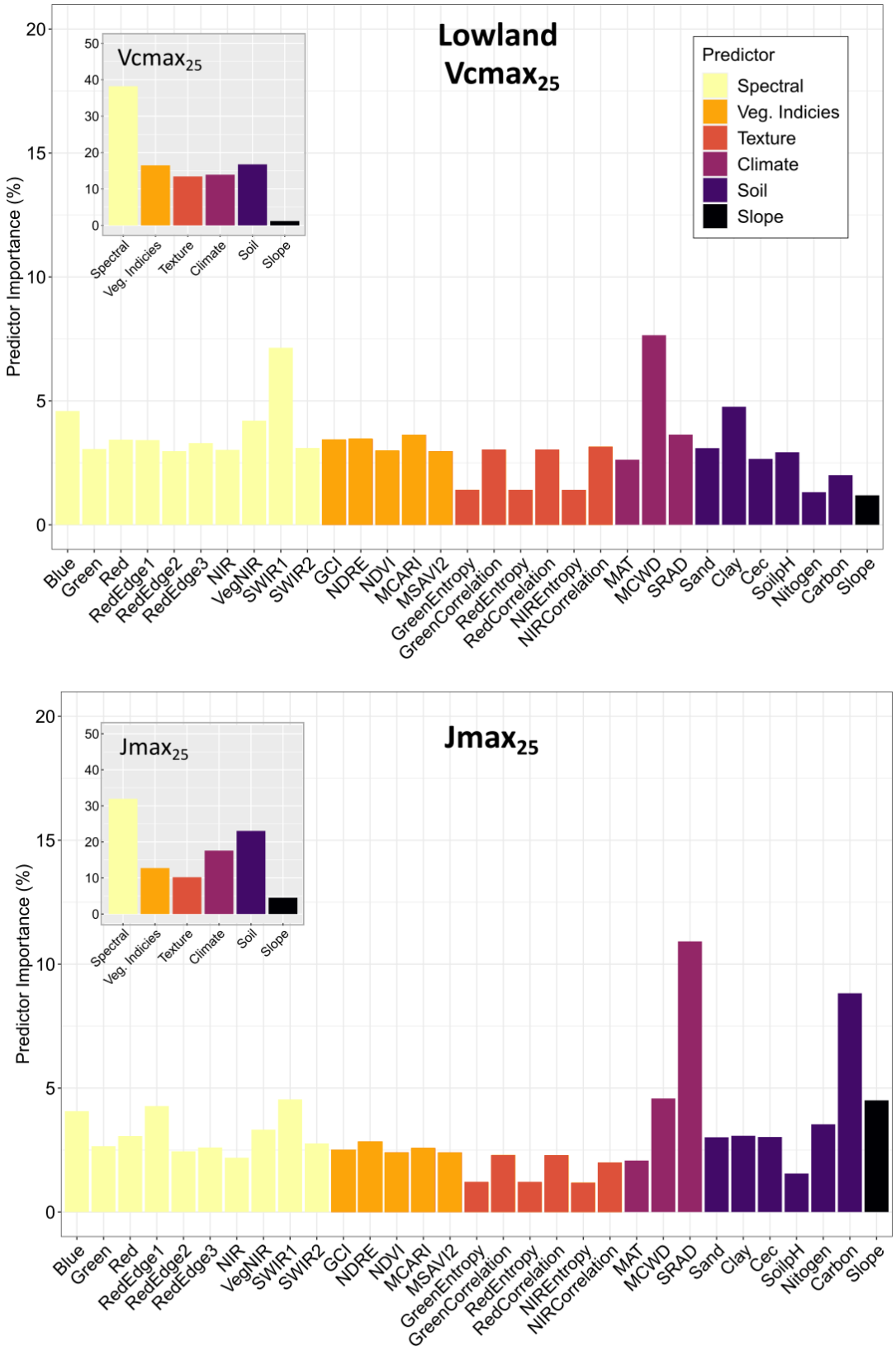


Figure S8.9: Percentage importance of each spectral, texture, climatic and soil variable for predicting lowland $Vcmax_{25}$ and $Jmax_{25}$. The same as **Figure 8.4**, but using single Sentinel images captured as close in time to photosynthesis sampling as possible instead of a 2019 Sentinel composite *Insert*: Total importance for each predictor group, represented as a percentage.

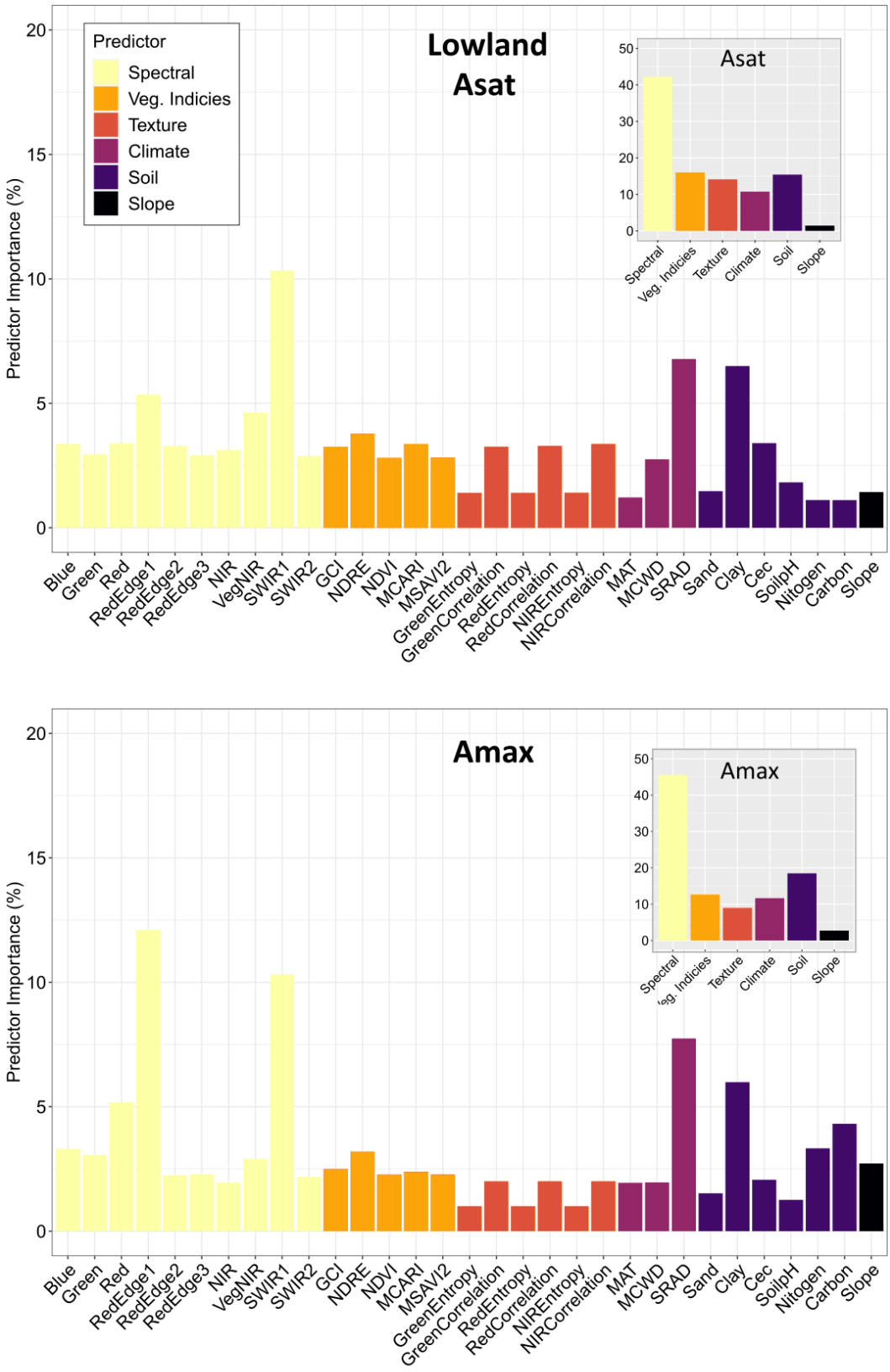


Figure S8.10: Percentage importance of each spectral, texture, climatic and soil variable for predicting montane Asat and Amax. The same as **Figure S8.7**, but using single Sentinel images captured as close in time to photosynthesis sampling as possible, instead of a 2019 Sentinel composite *Insert*: Total importance for each predictor group, represented as a percentage.

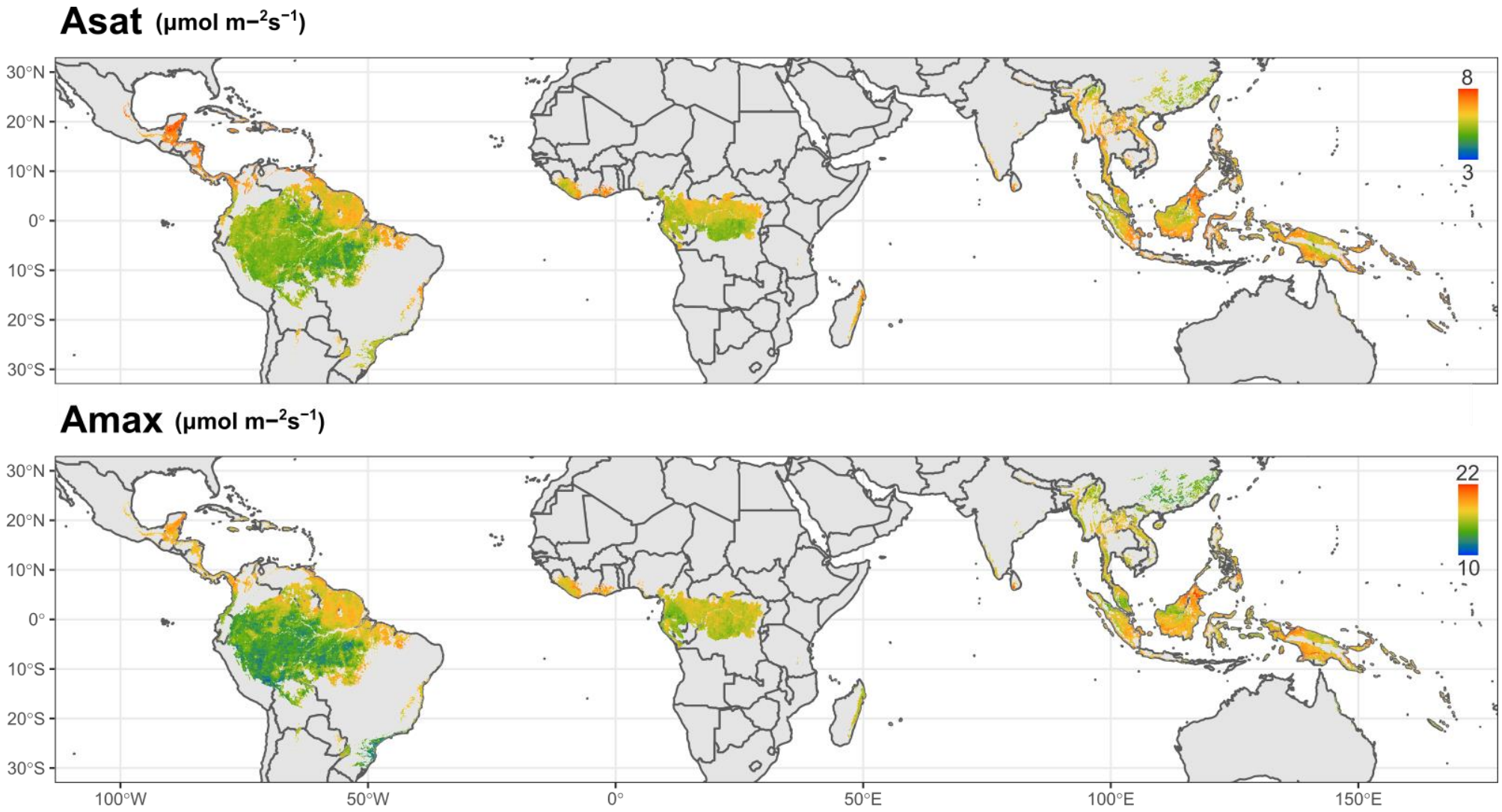


Figure S8.11: Pan-tropical maps of lowland Asat and Amax. Created in Google Earth Engine using field-collected trait values, Senitnel-2A imagery, environmental data and random forest regression. Original resolution at 10m. Resampled to 10km for display purposes. Only pixels located within the tropical forest biome, with >50% tree cover and <1000 m.a.sl are classified.

Publication	Location	Forest type	Soils	$Vcmax_{25}$ (µmol)	$Jmax_{25}$ (µmol)
Carswell *et al*. (2000)	Brazil	Wet evergreen	Nutrient-poor Oxisols or Ultisols	43 ± 6	97 ± 12
Domingues *et al*. (2007)	Brazil	Wet evergreen	Nutrient-poor Oxisols or Ultisols	52 ± 19	84 ± 32
Domingues *et al*. (2010)	West Africa	Semi-deciduous	Arenosols, Luvisols, Plinthosols, Alisols, Nitosols and Fluvisols	30 ± 10	50 ± 18
Domingues *et al*. (2010)	West Africa	Savannah	Arenosols, Luvisols, Plinthosols, Alisols, Nitosols and Fluvisols	71 ± 3	105 ± 17
Bahar *et al.* (2017)	Peru	Wet evergreen	Cambrisols, Alisols, Acrisols, Podzols and Arenosols	36 ± 15	67 ± 19
Bahar *et al.* (2017)	Peru	Tropical montane cloud forest	Cambrisols and Umbrisols	49 ± 20	97 ± 37
van de Weg *et al.* (2011)	Peru	Tropical montane cloud forest	Cambrisols and Umbrisols	56 ± 23	107 ± 47
Meir *et al*. (2007)	Cameroon	Wet evergreen	-	49–68	78–133
Domingues *et al*. (2015)	Cameroon	Forest-savannah transition	-	36–55	48–88
Kattge *et al*. (2009)	Global	Tropical forest	Nutrient-poor Oxisols	29 ± 8	-
Kattge *et al*. (2009)	Global	Tropical forest	Non-Oxisols	41 ± 15	-

Table S8.7: Tropical forest $Vcmax_{25}$ and $Jmax_{25}$ values found in the literature along with forest type and soil type (where available). $Vcmax_{25}$ and $Jmax_{25}$ from Domingues *et al.* (2007) and Meir *et al.* (2007) were standardized to 25°C by van de Weg *et al.* (2017).

Chapter 9:

Thesis discussion



9.1 Thesis synthesis

This doctoral thesis has examined the use of optical imagery for mapping vegetation properties at the leaf, UAV, aircraft and satellite level, through four research chapters. The main results are synthesised and compared below.

Chapter 5 evaluated the use of 5-band UAV imagery for mapping biochemical leaf traits in the High Arctic - an ecosystem where species richness and vegetation structure are minimised. Using field spectroscopy data (400-2500 nm) as a baseline, we could distinguish eight common High Arctic plant tundra species with 74% accuracy and predict their corresponding water, nitrogen, phosphorus and C:N values (r^2 = 0.61–0.88, RMSEmean = 12%–64%). At the UAV level we were unable to classify individual plant species, but mapped three plant functional types (mosses, graminoids and dwarf shrubs) at 72% accuracy, with the spectral separability of moss emerging as a significant result. Using UAV imagery to estimate biochemical plants traits resulted in lower results than the field spectroscopy data (r^2 = 0.44–0.54, RMSEmean = 26%–48%), even when the field spectroscopy data was downscaled to match the extent and spectral resolution of the UAV imagery. The reduction in accuracy was attributed to increases in imagery quality issues such as BDRF effects, varying illumination conditions during a flight, and spectral mixing across a heterogenous tundra.

Chapter 6 built on the results from **Chapter 5** and evaluated the use of hyperspectral UAV imagery (450-950 nm, 9nm FWHM) for mapping biochemical leaf traits across a tropical forest canopy in Ghana, West Africa. Similarly to **Chapter 5**, field spectroscopy data was used to generate a predictive baseline. Compared to **Chapter 5**, field-spectroscopy PLSR models were lower for similar leaf traits: r^2 = 0.42 (LMA), r^2 = 0.43 (N), r^2 = 0.21 (P), r^2 = 0.20 (K), r^2 = 0.23 (Ca) and r^2 = 0.14 (Mg). At the airborne level, four of the six leaf traits demonstrated weak (p < 0.10) correlations with the UAV-collected hyperspectral imagery: r^2 = 0.25 (LMA), r^2 = 0.22 (N), r^2 = 0.22 (P), and r^2 = 0.25 (Ca). Despite using higher spectral resolution UAV imagery than **Chapter 5**, the lower UAV results reported in **Chapter 6** are likely due to the higher species richness of a tropical forest ecosystem, and the increased structure and architectural properties of a tropical forest canopy that amplify BDRF effects through higher leaf area index (LAI), more

variable leaf angle distribution (LAD), and increased gap fraction and shadowing. In addition, the larger decreases in accuracy between the field spectroscopy and UAV models in **Chapter 6** compared to **Chapter 5**, may be attributed to asynchrony between the trait collection and the UAV imagery collection (52-64 weeks) and difficulty of matching the UAV imagery to the field-sampled tree crowns. However, ample signal was detected to map the leaf economic spectrum across a precipitation gradient, which reflected plot-level carbon cycling data, and revealed precipitation and substrate as co-dominant drivers of trait distributions and relationships.

Chapter 7 expanded a focus from mapping biochemical leaf traits to mapping process-based traits ($Vcmax_{25}$ and $Jmax_{25}$) across a forest disturbance gradient in Malaysian Borneo using aircraft-collected hyperspectral imagery, and Sentinel-2 imagery. Despite estimating traits not directly expressed in spectra, PLSR models for the aircraft-collected hyperspectral imagery were higher than those developed using the UAV-imagery in Ghana (r^2=0.36 and 0.43, RMSE= 20 and 31%, for $Vcmax_{25}$ and $Jmax_{25}$ respectively). An increase in model performance is likely due to the increased spectral extent (covering the SWIR, as well as the VNIR), tree crown maps measured to decimetre precision, which facilitated highly accurate matching of traits and spectral imagery, and a larger sample size (222 versus 58 tree crowns). Similarly to **Chapter 6**, field-measured traits and spectral imagery were measured with a temporal offset (32-56 weeks). **Chapter 7** also explored increased image quality issues associated with expanded spatial coverage. Aircraft-collected hyperspectral imagery was found to exhibit poor radiometric calibration and Sentinel-2 imagery was found to suffer from processing errors related to cloud masking and inconsistent surface reflectance values, especially over humid tropical forests. For Sentinel-2 imagery, strongest model results were produced using a time-series composite (r^2=0.21 and 0.28, RMSE= 22-34%) that outperformed a single Sentinel scene in predicting $Vcmax_{25}$ and $Jmax_{25}$ and also displayed the most credible reflectance patterns.

Chapter 8 built on the regional analysis of **Chapter 7** to create a pan-tropical map of $Vcmax_{25}$ and $Jmax_{25}$ by combining Sentinel-2 imagery, environmental datasets and photosynthesis data from 35 tropical forest GEM plots. Models captured 38% and 57% of variation in $Vcmax_{25}$ and $Jmax_{25}$ across lowland-only forest, and 62% and 61% of

variation across lowland-montane forest. These results were higher than those reported for predicting $Vcmax_{25}$ and $Jmax_{25}$ using Sentinel imagery in **Chapter 7**, likely due to the combination of Sentinel-2 imagery and environmental datasets, which harnessed the predictive power of environmental drivers of $Vcmax_{25}$ and $Jmax_{25}$ and reduced the influence of artefacts associated with Sentinel imagery. In lowland only forest, mean climatic water deficit (MCWD) and solar radiation were most important predictors of $Vcmax_{25}$ and $Jmax_{25}$ respectively. In lowland-montane forest, the strongest predictor of both $Vcmax_{25}$ and $Jmax_{25}$ was mean annual temperature (MAT). At intercontinental scales, $Vcmax_{25}$ and $Jmax_{25}$ were mainly controlled by climate. At the regional level, trait distributions were shaped by soil type. At local levels, trait patterns reflected spectral variation arising from topography and canopy structure.

Overall, this thesis offers new methodologies to measure and monitor tropical forest functional diversity and productivity. It builds from a simple ecosystem to a floristically and structurally complex tropical forest, and explores the trade-offs of expanding spatial coverage and decreasing spectral resolution for the retrieval of vegetation canopy traits. Alongside, it highlights current limitations of optical remote sensing, and presents best-practice recommendations for future imaging studies. The results are pertinent for high resolution, low cost mapping of tundra and tropical forest ecosystems, and for improving the representation of vegetation properties, especially photosynthetic capacity, in the next generation of Earth System Models. The main implications of these results and future research needs are discussed in section 9.2.

9.2 Main implications and future research needs

Synthesising the four research chapters, there are a number of theoretical and practical implications that can be drawn from their results to advance the optical remote sensing of vegetation canopy traits. This section discusses the main implications of these results and highlights future research needs.

Complementary trait and remote sensing campaigns

One of the major takeaways from this thesis is that the development of empirical models to estimate vegetation traits requires a large volume of field-sampled measurements. While empirical models combined with optical imagery can help scale up field measurements over the wider landscape, they are not a replacement for field surveys. In order to integrate field sampling and remote sensing imagery with the greatest accuracy, trait measurements must be standardized and spatially explicit. Currently, however, trait campaigns and remote sensing imagery acquisition are often planned separately, such as in **Chapters 6, 7** and **8**, resulting in spatial and temporal mismatches. For example, very small plots, or transect sampling are not well-matched to the size and shape of airborne and satellite pixels. In this thesis, we find that one-hectare plots are a good shape and resolution to match with UAV, aircraft and satellite imagery. Furthermore, we achieve significant results using a basal-area weighted sampling approach that records the trait values largest and most dominant species in a plot (Chapter **6, 7**, and **8**). In addition to sampling these species, we would also recommend sampling the tallest and largest individuals of these species (such as in **Chapter 7**), that are easily identifiable in optical imagery and contribute the most to canopy reflectance.

Temporally, trait collections are often planned as the main research objective, with remote sensing imagery acquisition organised as a later addition, or entirely separate study. Organising separate studies may lead to a temporal offset between trait values at the time of sampling and trait values at the time of imagery acquisition, resulting in lower accuracy models. A temporal offset may explain the weaker models reported in **Chapter 5** compared to **Chapter 6**. Across tropical forests, trait distributions vary seasonally and inter-annually: Evergreen tropical forests have been found to display

differences in trait values between dry and wet seasons (Gotsch *et al*. 2010; Asefa *et al*. 2021), and El Nino oscillations have been linked to inter-annual variability in trait distributions (Nunes *et al*. 2019). Planning complementary trait and remote sensing campaigns can help to reduce many of these spatial and temporal offset factors. Close collaboration among field scientists, airborne survey teams and stakeholders is required to design sampling methods that are suitable for all parties. While integrated sampling methods may often differ, or exceed the scope of tradition field surveys, they are required to advance the remote sensing of vegetation canopy traits and expand spatially limited plot studies.

Geolocating field measurements

A second major takeaway from this thesis when designing future airborne imaging studies, is the challenge of accurately geolocating field-sample measurements across airborne imagery. Accurately geolocating field-sample measurements was a challenge throughout all the chapters, but especially in **Chapter 5** and **Chapter 6.** In **Chapter 5,** the reduction in accuracy using UAV imagery compared to downscaled field spectroscopy data, was partly attributed to spatial offsets between the imagery and dGPS locations of turf samples. Even when imagery is georeferenced to decimetre precision, as in **Chapter 5**, small offsets between field-sampled vegetation areas and imagery across an extremely heterogenous ecosystem may result in inaccuracies between calibration and validation points. In these areas, we recommend covering nearby vegetation to create an exclusion zone around the sampled area, or a similar identification method.

Across tropical forests, due to the size of most tree crowns, small spatial offsets are less significant. However, GPS devices are much less accurate under a tropical forest canopy (accuracy >10cm) and do not record which crowns are fully sunlit, and which crowns are partially or fully-shaded. In **Chapter 6**, tree crowns had to be manually identified in spectral imagery using *xy* coordinates and expert species recognition, which was a labour-intensive task and resulted in a small sample size. Laser tree crown maps used in **Chapter 7** and **8** significantly improved the speed and accuracy of trait-spectra matching and may partially explain the stronger models reported in **Chapter 7** compared to **Chapter 6.** In our experience, geolocating field measurements across imagery is not

trivial and constituents one of the largest sinks of time and sources of uncertainty in this field. Improving geolocation methods, would significantly enhance the speed and accuracy of trait-spectra matching and likely lead to improved models.

Upscaling between field sampling and pixel size

A third important, but under-researched, area in optical remote sensing studies is the practice of 'upscaling'; extrapolation from fine-scale data to a coarser scale. Due to a common mismatch between field-sampled resolution and imagery pixel size, upscaling often takes place at the field-level. In **Chapter 5**, three 5 × 5 cm samples were used to represent the trait value of a 20 × 20 cm turf, and in **Chapters 6,7** and **8**, 3-5 leaf samples were used to represent the trait value of a tree crown. While our results showed that significant predictive signal was captured by these methods of upscaling, research is still lacking to show how much signal is lost through unrepresentative field data collection, and whether a more comprehensive, or adjusted, sampling method would lead to stronger predictive models. Studies have shown that trait values vary horizontally across tree crowns (Evans, 1979; Erdmann, 1988), and that leaf age, leaf health, and the presence of lianas affect canopy trait values and spectral reflectance (Sánchez-Azofeifa *et al.* 2009; Asner and Martin, 2012; Chavana-Bryant *et al.* 2019; Cotrozzi, 2021; Jiang *et al*. 2021). Although traits from degraded turfs were collected in **Chapter 5**, and trees with documented liana coverage in Brazil-ST were excluded from the analysis in **Chapter 8;** in general, testing methods of upscaling fell outside the scope of this thesis. Research exploring the most appropriate field sampling and scaling methods to maximise trait signal would likely lead to the development of stronger models, and may involve sampling all types of vegetation present in the study area, not just mature and healthy vegetation.

Image acquisition quality

In addition to improving the integration of field surveys and remote sensing imagery, this thesis has also shown that it important to improve the acquisition and processing quality of remote sensing imagery itself. **Chapter 7** presented a comprehensive evaluation of the strengths and weaknesses of different types of remote sensing imagery acquisition and processing. One of the main conclusions from **Chapter 7** was

that enhanced radiometric calibration is crucial to improve the accuracy of aircraft-collected imagery. At the satellite level, improved methods of cloud detection and masking would significantly increase the accuracy of Sentinel-2 imagery, along with improved aerosol optical depth and water vapour measurements for better surface reflectance retrieval. In terms of cloud masking, users must currently mask clouds themselves in Sentinel-2 imagery, using ESA's scene classification layer, which is around 75% accurate (Zekoll *et al*. 2021). Integrating a >90% accurate cloudless Sentinel product (e.g. Fmask4: Sanchez *et al*. 2020) into ESA's Sentinel Hub, or Google Earth Engine, would significantly increase the speed and performance of Sentinel-based mapping, and could be updated with the latest cloud-masking technology, as it become available.

Spatial resolution of environmental datasets

In **Chapter 8**, our results showed that combining Sentinel-2 imagery and environmental data layers significantly improved the performance of predictive models. However, many environmental data layers are in their infancy, such as the ISRIC SoilGrids, that represent the first attempt to map global soils at 250m and were first released in 2017, with several improvements added since then (Poggio *et al.* 2021). This mismatch in scale between current environmental data layers and plot-level data may be partly the reason why local models in **Chapter 8** failed to show significant correlations between photosynthetic capacity and spectral and environmental predictors. Higher resolution climatic and soil datasets, that captured differences between nearby plots, and especially datasets that captured microclimates and microtopography within plots, would likely significantly increase modelling accuracy.

Hyperspectral satellites

Lastly, our results show that the next generation of hyperspectral satellites (e.g. CHIME, EnMAP, HISUI, HypXIM, HyspIRI and PRISMA) will be instrumental in advancing the retrieval of vegetation properties globally. In **Chapter 5** and **6**, all field spectroscopy results were significantly higher than results achieved using lower spectral resolution data. In **Chapter 7**, stronger models were reported using aircraft-collected hyperspectral aircraft imagery compared with using multispectral Sentinel-2 imagery. Planned

hyperspectral satellites contain over x20 the spectral information currently captured by Sentinel-2 and have the potential to revolutionize the remote retrieval of vegetation properties. However, despite the clear advantage of increased spectral resolution, the results of this thesis show that many of the current limitations to the optical remote sensing of vegetation properties are not related to spectral resolution. There are still a number of areas where directed research is required to optimally harness current remote sensor technology, and prepare for future instruments. Designing complementary trait and remote sensing campaigns; accurately geolocating field measurements; testing representative upscaling methods; improving optical imagery quality; and increasing the spatial resolution and accuracy of environmental data layers are just a few of the areas where improvement could be made to significantly advance the mapping and monitoring of vegetation canopy properties in the one of the greatest eras of environmental change.

9.3. References

Asefa, M., Song, X., Cao, M., Lasky, J.R. and Yang, J. (2021) Temporal trait plasticity predicts the growth of tropical trees. *Journal of Vegetation Science,* 32 (4), pp. e13056.

Asner, G.P. and Martin, R.E. (2012) Contrasting leaf chemical traits in tropical lianas and trees: implications for future forest composition. *Ecology Letters,* 15 (9), pp. 1001-1007.

Chavana-Bryant, C., Malhi, Y., Anastasiou, A., Enquist, B.J., Cosio, E.G., Keenan, T.F. and Gerard, F.F. (2019) Leaf age effects on the spectral predictability of leaf traits in Amazonian canopy trees. *Science of the Total Environment,* 666, pp. 1301-1315.

Cotrozzi, L. (2021) Spectroscopic detection of forest diseases: a review (1970–2020). *Journal of Forestry Research,* , pp. 1-18.

Erdmann, G.G., Crow, T.R. and Rauscher, H.M. (1988) Foliar nutrient variation and sampling intensity for Acer rubrum trees. *Canadian Journal of Forest Research,* 18 (1), pp. 134-139.

Evans, J. (1979) The effects of leaf position and leaf age in foliar analysis of Gmelina arborea. *Plant and Soil,* 52 (4), pp. 547-552.

Gotsch, S.G., Powers, J.S. and Lerdau, M.T. (2010) Leaf traits and water relations of 12 evergreen species in Costa Rican wet and dry forests: patterns of intra-specific variation across forests and seasons. *Plant Ecology,* 211 (1), pp. 133-146.

Jiang, X., Zhen, J., Miao, J., Zhao, D., Wang, J. and Jia, S. (2021) Assessing mangrove leaf traits under different pest and disease severity with hyperspectral imaging spectroscopy. *Ecological Indicators,* 129, pp. 107901.

Nunes, M.H., Both, S., Bongalov, B., Brelsford, C., Khoury, S., Burslem, D.F.R.P., Philipson, C., Majalap, N., Riutta, T. and Coomes, D.A. (2019) Changes in leaf functional traits of rainforest canopy trees associated with an El Niño event in Borneo. *Environmental Research Letters,* 14 (8), pp. 085005.

Poggio, L., de Sousa, L.,M., Batjes, N.H., Heuvelink, G., Kempen, B., Ribeiro, E. and Rossiter, D. (2021) SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *Soil,* 7 (1), pp. 217-240.

Sanchez, A.H., Picoli, M.C.A., Camara, G., Andrade, P.R., Chaves, M.E.D., Lechler, S., Soares, A.R., Marujo, R.F.B., Simões, R.E.,O. and Ferreira, K.R. (2020) Comparison of Cloud cover detection algorithms on sentinel–2 images of the amazon tropical forest. *Remote Sensing,* 12 (8), pp. 1284.

Sánchez-Azofeifa, G.A., Castro, K., Wright, S.J., Gamon, J., Kalacska, M., Rivard, B., Schnitzer, S.A. and Feng, J.L. (2009) Differences in leaf traits, leaf internal structure, and spectral reflectance between two communities of lianas and trees: Implications for remote sensing in tropical environments. *Remote Sensing of Environment,* 113 (10), pp. 2076-2088.

Zekoll, V., Main-Knorn, M., Louis, J., Frantz, D., Richter, R. and Pflug, B. (2021) Comparison of masking algorithms for sentinel-2 imagery. *Remote Sensing,* 13 (1), pp. 137.