

Comparing the accuracy of remote sensing methods for estimating foliar pigments across tree species and years using DESIIS imagery

JESSICA SOTO BALVANERA

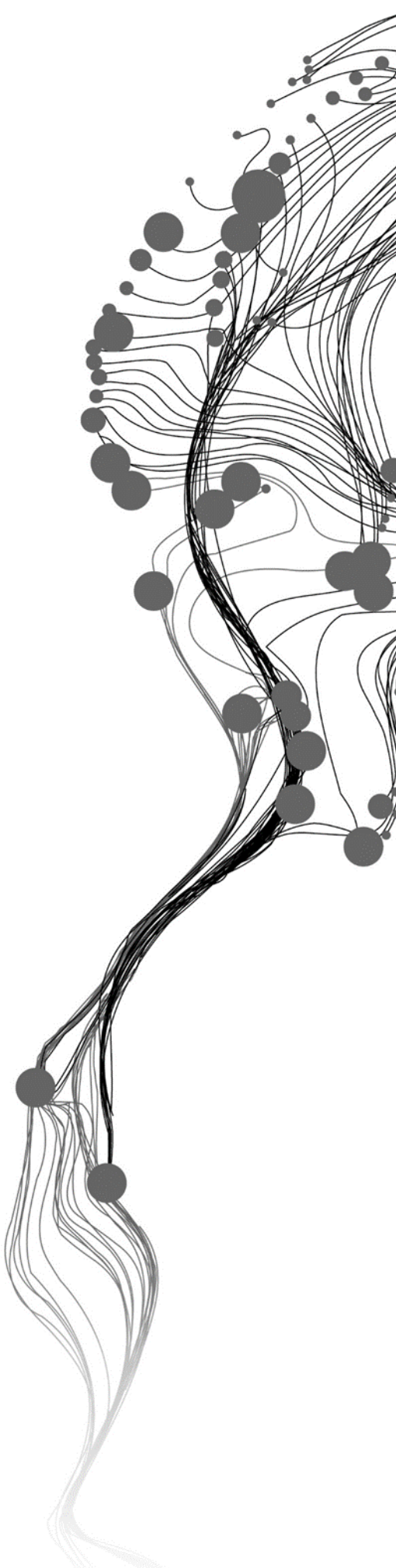
November, 2023

SUPERVISORS:

Dr. R. Darvishzadeh

Prof. Dr. A.K. Skidmore

A. Torres Rodriguez MSc (Advisor)



Comparing the accuracy of remote sensing methods for estimating foliar pigments across tree species and years using DESIIS imagery

JESSICA SOTO BALVANERA

Enschede, The Netherlands, November, 2023

Thesis submitted to the Faculty of Geo-Information Science and Earth
Observation of the University of Twente in partial fulfilment of the
requirements for the degree of Master of Science in Geo-information Science
and Earth Observation.

Specialization: Natural resources management

SUPERVISORS:

Dr. R. Darvishzadeh

Prof. Dr. A. K. Skidmore

A. Torres Rodriguez MSc (Advisor)

THESIS ASSESSMENT BOARD:

Dr. Ir. A. Vrieling (Chair)

T.W. Gara Ph.D (External Examiner, Cal Poly Humboldt)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

This study addresses gaps in the robustness of traditional models under diverse conditions and newly released high spectral resolution satellite imagery. It conducts a comparative analysis of established statistical methods, including partial least squares regression (PLSR) and narrowband vegetation indices (VIs), on DESIS hyperspectral satellite data to assess their adaptability across various tree species over a two-year period. A significant finding of the study is the varying accuracy of these methods across species, specifically between deciduous species and conifers. The research highlights the importance of the 500-510 nm green and 667.9-688.4 nm red spectral regions for deciduous species, alongside the 645-655 nm red spectrum and NIR bands at 827.2 nm and 927.9 nm for conifers. PLSR demonstrated good accuracy in estimating canopy chlorophylls and carotenoids in deciduous species (e.g., $R^2 = 0.66$ and $RMSE = 0.40$ for chlorophylls content in 2020), but this accuracy varied across species and years. Conversely, indices like the narrowband Datt derivative index (nDD) provided consistency but lacked accuracy. Employing VIs, a strong correlation was found between the 680-780 nm red edge region and the foliar pigments contents, consistent across years and species. The findings of this study emphasize the complexities involved in modelling canopy pigment content, emphasizing the necessity for species-specific methodologies and the integration of multi-temporal data in statistical models, to improve the retrieval accuracy of canopy pigments content and allow the accurate monitoring of stress and disturbance in forest ecosystems in an era of climate change and biodiversity loss.

Keywords: DESIS; hyperspectral imagery; canopy chlorophylls content; canopy carotenoids content; vegetation indices; partial least squares regression; model robustness; model consistency; forest monitoring

ACKNOWLEDGEMENTS

This MSc Thesis has been a journey of academic learning, personal development, and life experiences.

I am deeply grateful to my supervisors, Dr. Roshanak Darvishzadeh and Prof. Andrew Skidmore, for their trust and guidance throughout this process, and for granting me access to the BIOSPACE program datasets. I want to express my profound gratitude to my advisor, Alejandra Torres, whose work has been indispensable for the in-situ and laboratory measurements. Her invaluable support in sharing with me details and information about the data was only surpassed by her kindness and professionalism. Special thanks go also to Haidi Abdullah for his help with data processing, and to Tomas Groen for introducing me to the R programming language and his teachings in statistics. My heartfelt thanks to Stefani Holzwarth, who provided immense trust, support, and learning during my internship. I am forever grateful to ITC for the opportunity to study at this remarkable institution.

A strong support group was essential in completing my work. I am thankful to my friends, those who were always there for me in times when I needed to relax, to feel like home while away from my family or I needed a hand in difficult times. Thanks to the Frogs for welcoming me into their group and being like family. To the Crows for being such a fun and caring group. To Ama for being a great friend and always being there for me. To Gian for being my company and support at the final stage of my studies. To Enzo for being an amazing friend, who always had the patience to help me with my doubts and was up for playing overcooked or chess. To Lukas who kept reassuring me I would graduate and was always up for jumping in the Eisbach.

Eternas gracias a mi familia en México que me apoyó con esta locura de estudiar una maestría en Europa. Agradezco especialmente a mi mamá, quien siempre confió en que iba a poder lograrlo, incluso cuando todo parecía difícil, y a mi papá, quien me enseñó la importancia de la educación. A mis hermanas que me ayudaron sin pensarlo, en medio de una pandemia. A mi abuelita por apoyarme con las clases de inglés y quererme tanto. A la Animal, que cuenta como familia, por ser la mejor amiga del mundo y siempre confiar en mi cuando yo no puedo. A la familia Balvanera Cuevas, mi gratitud por su inmenso apoyo y confianza. A mis amigos, Karla, Zara, Estefanía y Arturo por su amistad y cariño, a los Rumberos por ser los más divertidos y bailadores. A Liza y María por ser la mejor compañía tanto para divertirnos como para trabajar.

The research was funded through the BIOSPACE project, financially supported by the European Research Council (ERC) within the framework of the European Union's Horizon 2020 research and innovation programme, as indicated by grant agreement number 834709. The authors would like to express their sincere appreciation for the support received during the field data collection process from the Bavarian Forest National Park (BNFP), the Bavarian State Forest Enterprise (Bayerische Staatsforsten), and Neuburg Forest in Germany. We are grateful for the support of the Hoge Veluwe National Park, Veluwezoom National Park and Het Loo Royal Estate in the Netherlands.

Finally, I extend my deepest gratitude to all those people who have been part of this incredible journey. Your support, in countless ways, has not only made this thesis possible but has also enriched my life beyond measure. This achievement is as much yours as it is mine.

TABLE OF CONTENTS

1.	Introduction.....	7
1.1.	Background.....	7
1.2.	Research gap.....	9
1.3.	Research objectives and questions	10
2.	Study sites and datasets.....	10
2.1.	Bavarian forest	10
2.2.	Veluwe forest	11
2.3.	Data collection and processing	12
2.4.	Ethical considerations.....	16
2.5.	Data preparation and pre-processing.....	17
3.	Methods.....	19
3.1.	Research methodology	19
3.2.	Narrowband vegetation indices	19
3.3.	Partial least squares regression	21
3.4.	Validation	22
4.	Results.....	23
4.1.	Narrowband vegetation indices	23
4.2.	Most relevant spectral regions.....	26
4.3.	Partial least squares regression	29
4.4.	Model performance comparison	30
5.	Discussion.....	31
5.1.	Accuracy of models.....	31
5.2.	Most important spectral regions for the estimation of canopy chlorophyll and carotenoids content	33
5.3.	Possible reasons for discrepancies in results and limitations	34
6.	Conclusions.....	35

LIST OF FIGURES

Figure 1. Location of the Bavarian Forest National Park. Base map source: Google Satellite	11
Figure 2. Location of Hoge Veluwe and Veluwe Zoom. Base map source: Google Satellite	12
Figure 3. Boxplot demonstrating the variation of canopy chlorophyll content in collected temperate species for the years 2020 and 2021.	15
Figure 4. Boxplot demonstrating the variation of carotenoid content in collected temperate species for the years 2020 and 2021.	15
Figure 5. Boxplot demonstrating the variation of carotenoid content in collected temperate species for the years 2020 and 2021.	16
Figure 6. Spectral curves of field measurements of deciduous plots in 2021 without smoothing.	18
Figure 7. Smoothed spectral curves of field measurements of deciduous plots in 2021 using the Savitzky-Golay filter.	18
Figure 8. The methodological framework of the study.	19
Figure 9. Year-specific 2D correlation plot illustrating the R^2 values between canopy chlorophylls (Cab) content and the narrowband Datt derivative index (nDD) calculated from each spectral band combination for (a) 2020 and (b) 2021.	27
Figure 10. Year-specific 2D correlation plot illustrating R^2 values between canopy carotenoids (Car) content and the narrowband Datt derivative index (nDD) calculated from each spectral band combination for (a) 2020 and (b) 2021.	27
Figure 11. Species-specific 2D correlation plot illustrating the R^2 values between canopy chlorophyll content (Cab) and the narrowband Datt derivative index (nDD) calculated from each spectral band combination for distinct tree species, (a) English oak, (b) European beech, (c) Norway spruce, and (d) Scots pine.	28
Figure 12. Species-specific 2D correlation plot illustrating R^2 values between canopy carotenoids content (Car) and the narrowband Datt derivative index (nDD) calculated from each spectral band combination for distinct tree species, (a) English oak, (b) European beech, (c) Norway spruce, and (d) Scots pine.	29
Figure 13. Panels (a) to (c) display scatter plots comparing in-situ measured canopy chlorophylls values with predictions from three different methods for the year 2020, panel (d) features a boxplot of both predicted and measured canopy chlorophylls values. The dashed lines in panels (a) to (c) indicate the ideal 1:1 correlation between predictions and in-situ measurements.	30

LIST OF TABLES

Table 1. Specifications of hyperspectral datasets.....	13
Table 2. Timeline of the 2020-2021 field campaign period.....	13
Table 3. Summary of the number of plots with specific canopy dominance in the 2020 and 2021 campaigns.....	13
Table 4. Descriptive statistics of field measurements for the years 2020 and 2021, for canopy chlorophyll a and b content (Cab) and carotenoids content (Car).....	14
Table 5. Descriptive statistics of field measurements for the distinct temperate species, for canopy chlorophyll a and b content (Cab) and carotenoid content (Car).....	15
Table 6. Selected narrowband vegetation indices and their respective formulations.	20
Table 7. RMSE and R ² of best band combinations using different vegetation indices equation for canopy chlorophylls data across years.....	23
Table 8. RMSE and R ² of best band combinations using different vegetation indices equation for canopy carotenoids data across years.	24
Table 9. RMSE and R ² of best band combinations using different vegetation indices equation for canopy chlorophylls data for different species.	25
Table 10. RMSE and R ² of best band combinations using different vegetation indices equation for canopy carotenoids data for different species.....	26
Table 11. RMSE and R ² obtained from PLSR for the estimation of canopy chlorophylls (Cab) and carotenoids (Car) data across two years.	29
Table 12. RMSE and R ² obtained from PLSR for the estimation of canopy chlorophylls (Cab) and carotenoids (Car) data across different species.....	29

1. INTRODUCTION

1.1. Background

Natural and anthropogenic stressors, including diseases, soil nutrient imbalances, water shortages or excesses, interactions with other species (e.g., pests and herbivores), and pollution, can have a profound impact on pigment content (Lichtenthaler, 1998; Sanchez et al., 1983). These factors make variations in chlorophyll a + b (Cab) and carotenoids (Car) content valuable indicators of physiological changes (Mohammed et al., 2000; Sims & Gamon, 2002). During conditions like plant stress and senescence, pigment concentrations often change, resulting in changes in the carotenoid-to-chlorophyll ratio: carotenoid content tends to increase, while chlorophyll content decreases (Baltzer & Thomas, 2005; Merzlyak et al., 1999; Peñuelas & Filella, 1998). Moreover, several factors, including chlorophyll and carotenoid content, water content, leaf structure, and the presence of proteins, influence the photosynthetically active absorbed radiation in the plant canopy (Gitelson et al., 2015; Homolová et al., 2013). Accurately estimating the temporal and spatial distribution of foliar pigment concentrations is of utmost importance for forest managers, ecologists, and farmers. It facilitates the monitoring of vegetation canopy adaptation, physiology, and functionality, especially in the face of critical challenges including biodiversity loss, climate change, and food security (Féret et al., 2017). Appropriately, it has been recognized as an Essential Biodiversity Variable (EBV), contributing to the assessment of the Aichi Biodiversity Targets (Skidmore et al., 2015).

Chlorophylls serve as the primary plant pigments responsible for the process of photosynthesis. They encompass chlorophyll a and b, collectively referred to as 'total chlorophyll'. The core physiological role of chlorophyll involves the absorption of incident radiation and the subsequent conversion of this energy into chemical energy within the photosynthetic apparatus. The concentration of foliar chlorophyll is indicative of the overall health of ecosystems and plays a pivotal role in determining the potential for photosynthesis and, consequently, primary production (Gholz et al., 1997). In addition to chlorophyll, the second major group of foliar pigments is carotenoids (abbreviated as Car). Carotenoids encompass carotenes and xanthophylls. Their primary physiological function lies in light absorption, but they also serve a secondary role in photoprotection (Ritz et al., 2000).

Traditional methods for measuring foliar pigments involve the extraction of pigments using solvents and subsequent spectrophotometric analysis of the resulting solutions, as described by Blackburn (2007). These methods allow researchers to estimate individual pigments based on the unique absorption properties of each pigment (Lichtenthaler, 1987). However, it is important to note that these conventional practices entail the collection of leaf samples, a process that can be destructive to the vegetation and ecosystems. Moreover, they necessitate laboratory equipment that can be both time-consuming and costly. These limitations hinder the ability to effectively monitor the temporal and spatial dynamics of vegetation health and stress (Sims & Gamon, 2002).

Remote sensing (RS) technologies, also frequently termed Earth Observation (EO) technologies, play a crucial role in assessing and monitoring vegetation across various temporal and spatial scales (Ali et al., 2020). While traditional multispectral spectroradiometers offer a limited number of broad spectral bands with gaps between them, making them unable to capture a continuous spectrum of canopy reflectance (Lu et al., 2019), hyperspectral data, characterized by hundreds of narrow bands often less than 10 nm in width, provides the detailed information needed for the analysis of biochemical and biophysical properties of vegetation (Lucieer et al., 2014).

The deployment of hyperspectral technology, particularly portable spectroradiometers, has facilitated non-intrusive research in the estimation of photosynthetic pigments (e.g., Datt, 1999; Maccioni et al., 2001; Zhang et al., 2021), and such techniques are commonly implemented in field studies. Subsequently, the utilization of airborne hyperspectral imagery became prevalent in scientific studies (e.g., Hoepfner et al., 2020; Sampson et al., 2003; Zarco-Tejada et al., 2001). It is worth noting that acquiring airborne hyperspectral imagery is a technically intricate and costly process that demands extensive planning prior to acquisition (Hruska et al., 2012).

Hyperspectral satellite sensors available in the past had restrictions in both spatial and temporal coverage (Lu et al., 2019). EO-1 Hyperion, in particular, was the sensor of choice for numerous studies employing hyperspectral satellite imagery (e.g., Wu et al., 2008, 2010; Yang et al., 2015a). Due to these inherent spatial and temporal limitations, the majority of past studies primarily focused on the assessment of chlorophyll and carotenoid content at the leaf level, predominantly relying on portable spectroradiometers (e.g., Jiang et al., 2022; Lichtenthaler et al., 1996; Sonobe & Wang, 2017). The recent improvements in hyperspectral - increasingly termed image spectroscopy - satellite technologies have led to significantly broader global coverage and enhanced spatial and temporal resolution, greatly improving the capability for non-destructively determining individual photosynthetic pigments.

This evolution in hyperspectral satellite technology includes several notable developments. The PRISMA (PRecursore IperSpettrale della Missione Applicativa) mission, developed by the Italian Space Agency (ASI) and launched in 2019, stands as a significant milestone. Additionally, sensors developed by the German Aerospace Center (DLR), including the Earth Sensing Imaging Spectrometer (DESI) launched in 2018 and the Environmental Mapping and Analysis Program (EnMAP) in 2022, contribute to the ongoing progress in this field. Future missions involving space-borne spectrometers like the Hyperspectral Infrared Imager (HypIRI) and the Copernicus Hyperspectral Imaging Mission for the Environment (CHIME) hold the potential for further advancing hyperspectral remote sensing applications. Specifically, DESI is a hyperspectral instrument designed and built by DLR, integrated into the Multi-User-System for Earth Sensing (MUSES) platform aboard the International Space Station (ISS). DLR's primary focus is to utilize the data acquired through DESI for scientific research within the realms of earth and atmospheric sciences, with the goal of enhancing our current understanding of passive hyperspectral remote sensing (Eckardt et al., 2015).

Various methods are available to quantify biochemical variables from remote sensing data. Verrelst et al. (2015) classified these methods into three main categories in a comprehensive review. First, statistical approaches encompass parametric and non-parametric techniques. Parametric methods assume predefined relationships between biochemical variables and spectral measurements, often utilizing band formulations like vegetation indices (VIs). Non-parametric methods derive linear and non-linear functions directly from the data, such as stepwise multiple linear regression (SMLR) and partial least squares regression (PLSR). Second, physical-based methods rely on established physical laws and cause-effect interactions, employing radiative transfer functions to acquire model variables. Examples include the PROSPECT and LEAFMOD models. Third, hybrid approaches combine elements from physical-based models and non-parametric statistics, offering flexibility in estimating biochemical variables. These diverse approaches provide a range of tools for remote sensing-based biochemical variable quantification.

Empirical relationships involving hyperspectral VIs are frequently utilized to estimate foliar pigments (Yang et al., 2015b). These VIs are derived from specific combinations of narrow spectral bands, making the identification of key bands crucial for practical applications. For instance, Chen et al. (2007) identified the significance of spectral regions near the green peak around 550nm and the red edge position (REP) for estimating pigments like chlorophyll a, b, and carotenoids in rice crops. Similarly, Hoepfner et al. (2020) explored the utility of airborne hyperspectral imagery in assessing forest canopy chlorophyll content, pinpointing important spectral regions in the visible range between 390-400 nm and 460-540 nm, in the near-infrared range around 1050-110 nm, and in the shortwave infrared region spanning 2000-2700 nm. In a prior study conducted in two broadleaf forests, Sonobe & Wang (2018) found that spectral regions around 500, 516–517 nm in the visible spectrum and between 744–750 nm in the near-infrared spectrum yielded the most reliable results for estimating carotenoid content.

1.2. Research gap

Up to now, most of the research in the field of vegetation pigments has concentrated on chlorophyll, with relatively limited attention directed toward the assessment of carotenoids at a canopy level (e.g., Blackburn, 1998; Hernández-Clemente et al., 2012, 2014; Kong et al., 2017; Miraglio et al., 2019, 2022; Sonobe & Wang, 2018; Stagakis et al., 2010; Yi et al., 2014; Zarco-Tejada et al., 2013). Monitoring the changes in diverse pigment contents, such as carotenoids, is essential to tackle several vegetation-related challenges, encompassing plant stress, diseases, invasive species, and photosynthetic phenology (Féret et al., 2017).

Previous research suggests that statistical approaches exhibit less generality than physically based models as they tend to be species, site, and time-specific due to their reliance on measured data (Ali et al., 2020; Zhen et al., 2021). However, the literature review reveals a gap, with no studies investigating the cross-species and temporal consistency of statistical methods for retrieving canopy chlorophyll (Cab) and carotenoid (Car) content using hyperspectral satellite imagery in forest ecosystems. Many of the reviewed studies are grounded in agricultural settings, comparing crop species (Croft et al., 2020; Ju et al., 2010) and consistency across years (Zhang et al., 2021), while others have employed airborne sensors (Hernandez-Clemente et al., 2014; Miraglio et al., 2019). The development of robust and operational methods for estimating foliar pigments at the canopy scale is critical for minimizing the reliance on model calibration, an expertise demanding process, field sampling, and high computational capacity.

1.3. Research objectives and questions

Study aim:

To assess how the accuracy of RS models for canopy chlorophylls and carotenoids content estimation varies among distinct tree species and across different years, using hyperspectral data.

Specific objectives:

1. To evaluate and compare the spectral bands/regions that are the most important for canopy chlorophylls and carotenoids content estimation in distinct tree species (Norway spruce, European beech, Scots pine, and English oak) and two successive years using hyperspectral satellite data.
2. Analyse and compare the accuracy and precision in predicting canopy chlorophylls and carotenoids content of multiple narrow-band vegetation indices (VIs) and partial least squares regression (PLSR) for distinct tree species and over two successive years.

Research questions

1. Which narrow spectral bands/regions are most important for the estimation of canopy chlorophylls and carotenoids content in the studied tree species and two successive years?
2. How does the accuracy of remote sensing statistical models (narrowband VIs and PLSR) vary when developed for specific tree species (Norway spruce, European beech, Scots pine, and English oak) and over two successive years?

2. STUDY SITES AND DATASETS

2.1. Bavarian forest

This study focuses on two temperate forest sites. The Bavarian Forest National Park (BFNP) is situated in southeast Germany (48°58'N, 13°23'E) beside the border with the Czech Republic. It covers an area of 24,369 hectares and forms part of the Eastern Bavarian Forest, which along with the Czech Bohemian Forest constitutes one of the largest Central European continuous forests (Heurich et al., 2010). The park's elevation ranges from 600 meters to 1453 meters, it has a temperate climate with mean annual temperatures from 3 to 6 °C and annual precipitation varies between 1200 mm and 1800 mm (Ali et al., 2020). The national park is composed of three main forest types (Sommer et al., 2015). Sub-alpine spruce forests are in the highlands (over 1,100 m a.s.l.) and cover 16 percent of the area with Norway spruce (*Picea abies*) as the dominant species. The slopes are characterized by mixed mountain forests (from 600 to 1,100 m a.s.l.) covering 68 percent of the BFNP, with Norway spruce, European beech (*Fagus sylvatica*) and silver fir (*Abies alba*) as prevailing species. Spruce forests cover the rest of the area, with Norway spruce, mountain ash (*Betula pendula*), and birch (*Betula pubescens*) prevailing in the valleys.

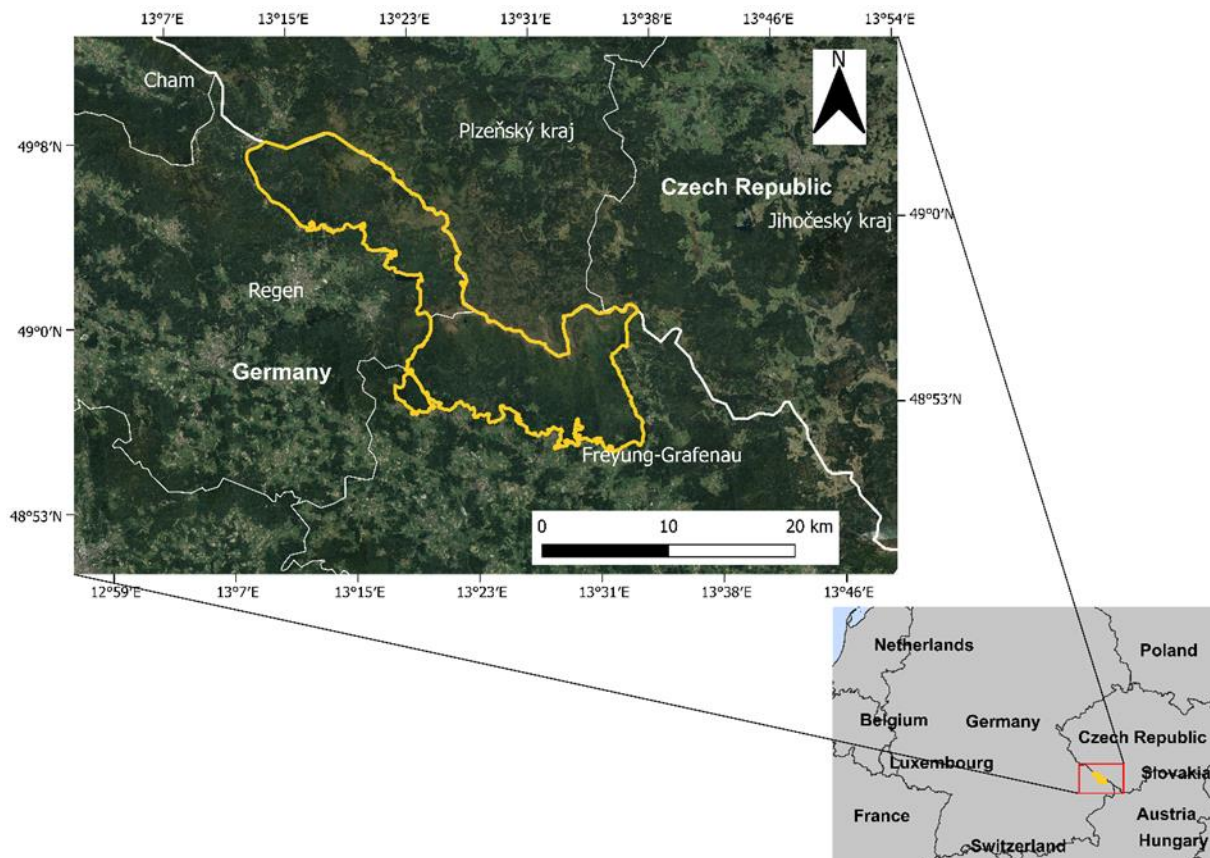


Figure 1. Location of the Bavarian Forest National Park. Base map source: Google satellite

2.2. Veluwe forest

The Veluwe forest (52°11'–52°24' N, 5°71'–5°92' E) is situated in the centre of the Netherlands, in the province of Gelderland. The climate of the Netherlands is temperate oceanic with a mean air temperature of 14.1 °C and mean annual precipitation of 832.5 mm (Neeffjes, 2018). By the 19th century the land cover changed from forest to open heathlands and drift sands by 50 percent due to overgrazing and deforestation, these areas deteriorated, and their agricultural use diminished, consequently, degraded zones were planted with pine and deciduous trees (Plakman et al., 2020). The Veluwe Zoom NP and the Hoge Veluwe NP are study areas within the Veluwe forest. The Veluwe Zoom NP (52°00'N, 06°01'E) is located northeast of the city of Arnhem, covering an area of approximately 5000 hectares. The Veluwe Zoom entails mostly deciduous forests mixed with farmlands and pastures (Mos et al., 2014). The Hoge Veluwe NP (52°06'N, 05°52'E), is located between the villages of Otterlo and Hoenderloo, it has an area of about 5500 hectares (Turnhout et al., 2004). Approximately 2282 hectares of coniferous forest and 613 hectares of deciduous forest cover the Hoge Veluwe National Park, along with heathlands and drift sands (Hein, 2011).

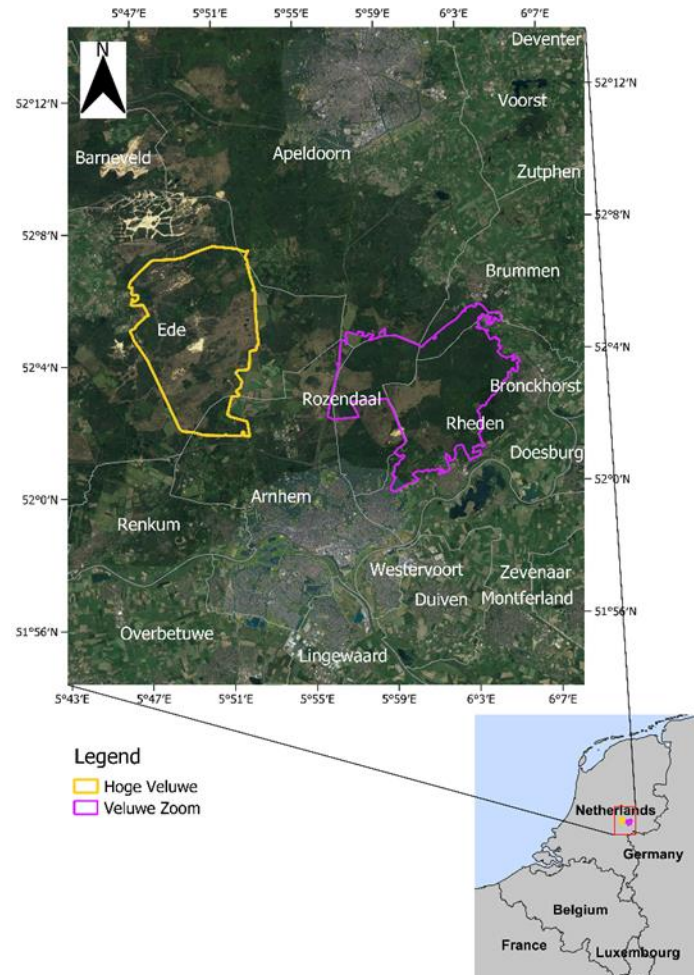


Figure 2. Location of Hoge Veluwe and Veluwe Zoom.
Base map source: Google satellite

2.3. Data collection and processing

The research presented in this thesis was a part of the BIOSPACE project, financially supported by the European Research Council (ERC) within the framework of the European Union's Horizon 2020 research and innovation programme, as indicated by grant agreement number 834709. The data employed consisted of secondary sources and can be categorized into two primary sections:

2.3.1. Hyperspectral

DESIS hyperspectral data were acquired for both study sites in the summer of 2020 and 2021 (Table 1). DESIS operates across a continuum of 235 spectral bands, spanning from 400 to 1000 nm in the visible near-infrared region, enabling a spectral sampling distance of 2.5 nm and a ground sampling distance (GSD) of 30m (Kruz et al., 2019). To obtain the canopy reflectance, the spectra from the DESIS imagery were extracted from the pixel corresponding to the centroid of the field plots (30x30m), this was done for each year and location. Out of a total of 220 plots that were surveyed, 48 plots were excluded from the study due to their location falling outside the DESIS image extent.

Table 1. Specifications of hyperspectral datasets.

Site	Date of acquisition	Sun zenith angle	Scene incidence angle	% Cloud shadow
Bavaria	01/07/2020	55.21	19.22	11.57
Veluwe	23/06/2020	45.54	3.80	1.34
Bavaria	17/06/2021	26.54	3.12	0.22
Veluwe	16/06/2021	29.77	4.11	1.25

2.3.2. Field measurements

In-situ measurements were collected during the summer of 2020 and 2021 (Table 2). Plots were randomly distributed within the study sites and had an area of 30 x 30 m. Two to three representative trees within each plot were selected for sampling. A Leica differential Global Positioning System (GPS) was employed to log the central coordinates of each plot. This system offers a geometric accuracy of less than a meter after post-processing. Sunlit leaves from the top of the canopy of the trees were recollected using a slingshot. The samples were stored in airtight plastic bags surrounded by ice blocks within an ice cooler for a maximum of six hours for transportation, the final storage was in a lab freezer. LAI was measured for each plot using the LAI 2200 Plant Canopy Analyzer (Li-COR). Five readings were taken below the canopy, within the plot boundaries, and three calibration readings (above the canopy) were taken at the nearest open area.

Table 2. Timeline of the 2020-2021 field campaign period.

Site	Year	Campaign period
Bavaria	2020	July - August
	2021	July - August
Veluwe	2020	June - July
	2021	May - June

A detailed description of the datasets for both the 2020 and 2021 campaigns is provided in Table 3.

Table 3. Summary of the number of plots with specific canopy dominance in the 2020 and 2021 campaigns.

Site	Year	European beech	Norway spruce	Scots pine	Mixed	English oak	Silver fir	Birch	Total
Bavaria	2020	16	13	0	15	0	1	0	45
Veluwe	2020	2	3	4	1	1	2	3	16
Bavaria	2021	16	16	0	0	0	0	0	32
Veluwe	2021	19	16	19	0	25	0	0	79
Total		53	48	23	16	26	3	3	172

2.3.3. Wet chemistry analysis of pigment content

The chlorophyll and carotenoids content in leaves were measured using two disks of different leaf sections on each tree sampled per plot. The extraction process involved adding 6 ml of acetone buffered with 10 g of MgCO₃, and Leaf disintegration was facilitated employing liquid nitrogen and sand. A UV/Vis spectrophotometer was used to measure the absorbance. The concentrations of chlorophyll a + b (Cab) and carotenoids (Car) in the solutions were determined by extracting the absorbance readings at specific wavelengths, following the equations by Lichtenthaler and Buschmann (2001), a widely recognized standard for measuring chlorophyll and carotenoid concentrations in plant samples. This method enabled the calculation of chlorophylls and carotenoids content per unit of leaf area (µg/cm²). For each tree, the measurements were done in duplicates (< 10% CV), and the resulting leaf chlorophylls and carotenoids content was averaged for a single measurement per plot. Subsequently, the collected measurements were multiplied by the corresponding Leaf Area Index (LAI) to upscale them to the canopy level, as outlined in Equations 1 and 2. These values were then converted into grams per square meter (g/m²), a unit conventionally utilized in canopy level studies.

$$\text{Canopy Cab} = \text{Leaf Cab} \times \text{LAI} \quad (1)$$

$$\text{Canopy Car} = \text{Leaf Car} \times \text{LAI} \quad (2)$$

Table 4. Descriptive statistics of field measurements for the years 2020 and 2021, for canopy chlorophyll a and b content (Cab) and carotenoids content (Car).

Statistics	2020		2021	
	Cab (g/m ²)	Car (g/m ²)	Cab (g/m ²)	Car (g/m ²)
Number of samples	57		111	
Minimum	0.53	0.09	0.13	0.04
Maximum	3.08	0.56	3.35	0.68
Range	2.55	0.47	3.21	0.65
Mean	1.91	0.33	1.29	0.25
Coefficient of variation	0.36	0.37	0.55	0.53
Standard deviation	0.69	0.12	0.71	0.13

Table 5. Descriptive statistics of field measurements for the distinct temperate species, for canopy chlorophyll a and b content (Cab) and carotenoid content (Car).

	Norway spruce		European beech		Scots pine		English oak	
Statistics	Cab (g/m ²)	Car (g/m ²)	Cab (g/m ²)	Car (g/m ²)	Cab (g/m ²)	Car (g/m ²)	Cab (g/m ²)	Car (g/m ²)
Number of samples	47		50		23		26	
Minimum	0.77	0.12	0.33	0.07	0.54	0.09	0.13	0.04
Maximum	2.98	0.54	3.35	0.68	1.17	0.24	1.81	0.28
Range	2.21	0.42	3.02	0.61	0.64	0.15	1.67	0.24
Mean	1.72	0.31	1.8	0.35	0.92	0.17	0.67	0.13
Coefficient of variation	0.33	0.34	0.43	0.38	0.2	0.21	0.63	0.48
Standard deviation	0.56	0.11	0.77	0.13	0.18	0.04	0.42	0.06

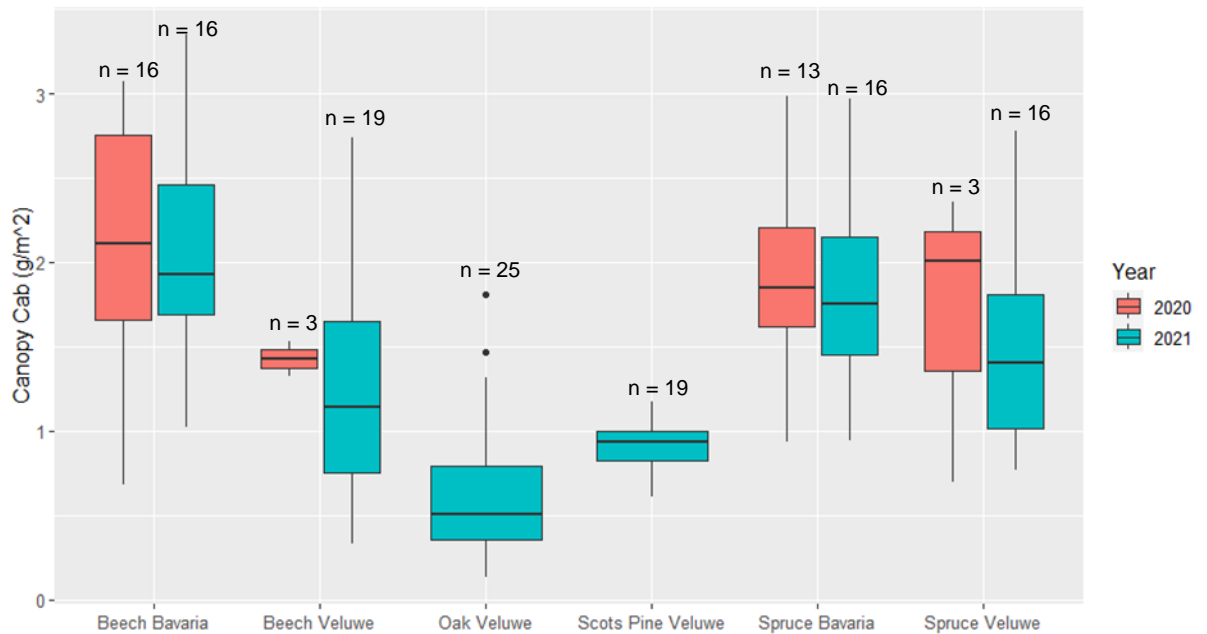


Figure 3. Boxplot demonstrating the variation of canopy chlorophyll content in collected temperate species for the years 2020 and 2021.

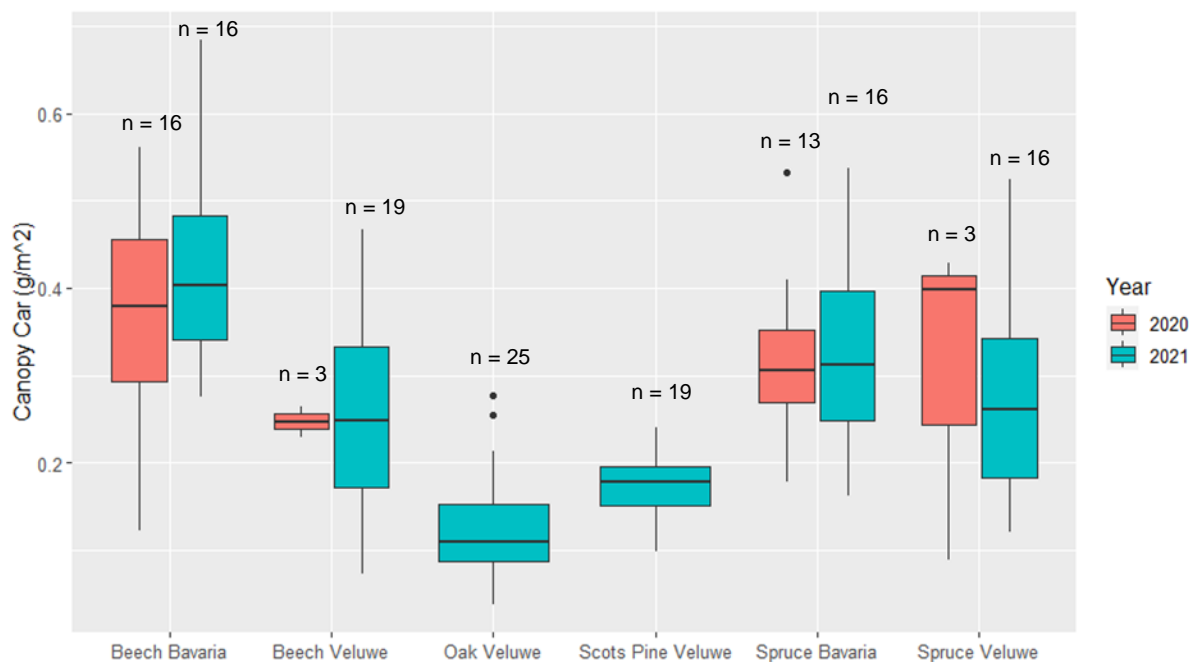


Figure 5. Boxplot demonstrating the variation of carotenoid content in collected temperate species for the years 2020 and 2021.

2.4. Ethical considerations

In this study, which employs secondary data, authorization was obtained to utilize field data and DESIS imagery, facilitated by the BIOSPACE program and the German Aerospace Center (DLR). A commitment was made to the intellectual property rights of DLR, and the professionals engaged in the collection and processing of this data. Our responsibility includes adhering to the terms specified in our agreements and being attentive to any future amendments. An important aspect of our research is its focus on the foliar pigment content of temperate forest ecosystems. Given the nature of this data, which is neither confidential nor sensitive, additional ethical considerations typically required for sensitive data were deemed unnecessary. Our approach ensures compliance with standard research ethics while recognizing the specific context of our data.

The aim of this research is to contribute to the broader scientific community. To this end, we intend to make our findings and as much data as possible available in an open-access format, in line with the consensus reached with all participating parties. Additionally, it is imperative for subsequent studies that build upon our conclusions to be cognizant of the limitations as outlined in our final document. Understanding these constraints is essential to ensure the continued validity and applicability of research in this field.

2.5. Data preparation and pre-processing

Pre-processing is an essential preliminary phase in the domain of hyperspectral imagery. A common practice involves eliminating problematic spectral bands, which encompasses addressing issues like atmospheric water absorption bands and low signal-to-noise ratio bands, since the inclusion of these bands can significantly impact the accuracy of predictive models (Ji et al., 2019). Vaiphasa (2006) states that signal noise is a concurrent problem in image spectroscopy, since the narrow bandwidth of hyperspectral sensors capture small amounts of energy that can be interfered by the noise generated inside the sensor. Also, physical external factors like variations in light illumination and atmospheric conditions further reduce the accuracy of the recorded spectral signals. Hyperspectral remote sensing studies frequently use spectral smoothing and aggregating techniques, to remove noise from spectral data (Y. H. Li et al., 2021). Atmospheric correction is a crucial initial step in handling challenges associated with hyperspectral images since, despite their information richness, these datasets often suffer from atmospheric distortions and may not faithfully represent the true surface reflectance (Rani et al., 2017). These steps ensure the imagery is optimized for clarity, precision, and consistency, making it suitable for further data manipulation. In this study, the DESIS datasets, being Level 2A products, underwent atmospheric correction through DLR's preprocessing chains.

2.5.1. Co-registration

Georeferencing is the process of associating images with specific ground-based coordinate systems. This process is essential for converting images from satellites, aerial perspectives, or ground-based sources into valuable mapping assets (Zhu et al., 2008). To ensure the precision of georeferencing for the DESIS datasets, they were cross-checked against Sentinel-2A imagery from the same location and date. The Image-to-image co-registration processing QGIS plugin was used to spatially align the target image (DESI) to the reference image (Sentinel-2A). An automated global algorithm was used to adjust the entire image on both distance and direction. Following this, the aligned image was set against its original to determine the extent and direction of the shift. This entire process was conducted for each individual image in the dataset.

2.5.2. Outlier and band removal

The hyperspectral datasets were thoroughly examined. A comparative analysis of sample reflectance was conducted, wherein each sample's spectral signature was visually compared with those of other samples. Within this analysis, it was observed that three European beech plots and one Norway spruce plot exhibited deviating reflectance characteristics, inconsistent with the typical spectral profile of healthy vegetation. These outlier plots were situated near non-forested pixels, thus raising concerns about potential georeferencing errors. Such misalignment can significantly impact the accuracy of reflectance values derived from remote sensing data, as the selected pixel may inadequately represent the plot's attributes. Consequently, the decision was made to categorize these four plots as outliers and exclude them from further analysis.

In the preprocessing of the DESIS hyperspectral dataset, a careful selection of spectral bands was performed to enhance the accuracy of chlorophyll content retrieval. To mitigate noise, and minimize redundancy, the first 10 bands were excluded from the analysis, due to low signal-to-noise when compared to the other 215 longer wavelength bands. The exclusion of these bands aligns with established practices in hyperspectral data processing, as the initial bands of a sensor are susceptible to more interference due to various factors such as atmospheric conditions and sensor limitations (X. Li et al., 2022).

2.5.3. Smoothing

The Savitzky-Golay filter, a low-pass filtering technique, is based on the premise that neighbouring data points exhibit significant similarities. This characteristic enables it to effectively mitigate noise, as mentioned by Flannery (2007). Its effectiveness in handling hyperspectral data has been highlighted in studies by Tsai & Philpot, 1998 and Schläpfer et al., 2011. To enhance the quality of the data, reflectance values from each plot were subjected to smoothing using the Savitzky-Golay filter, implemented through the 'sgolay' R package. The filter was configured with a window size of seven and a second-degree polynomial order. The reflectance for deciduous trees in 2021 can be seen in Figure 5 and the smoothed spectra in Figure 6.

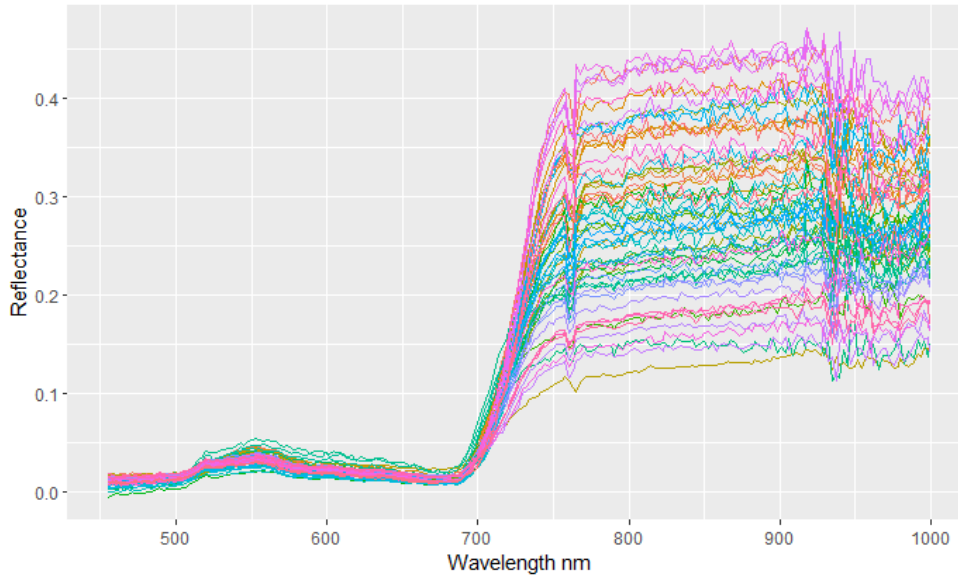


Figure 6. Spectral curves of field measurements of deciduous plots in 2021 without smoothing.

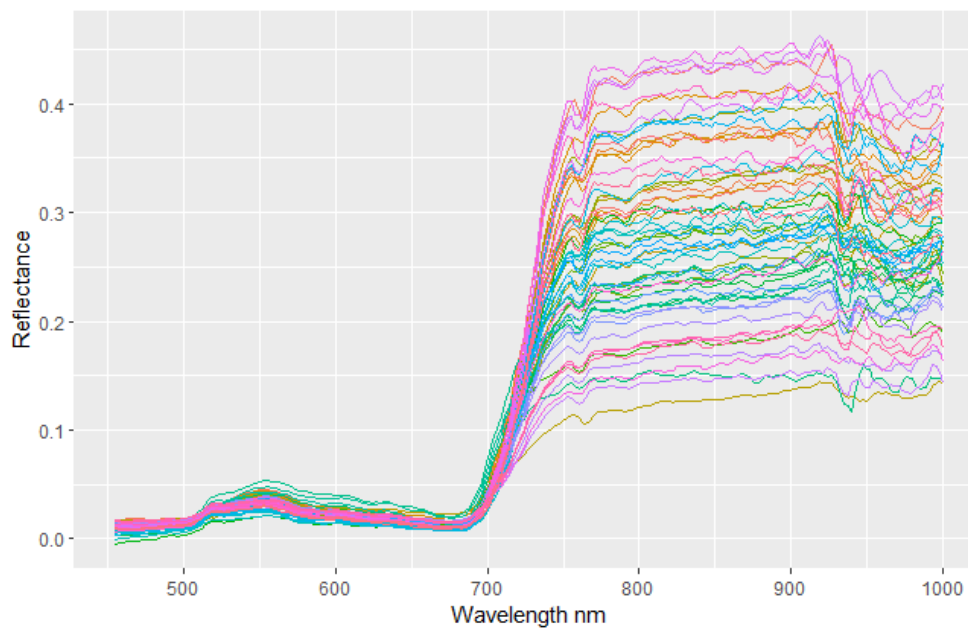


Figure 7. Smoothed spectral curves of field measurements of deciduous plots in 2021 using the Savitzky-Golay filter.

3. METHODS

3.1. Research methodology

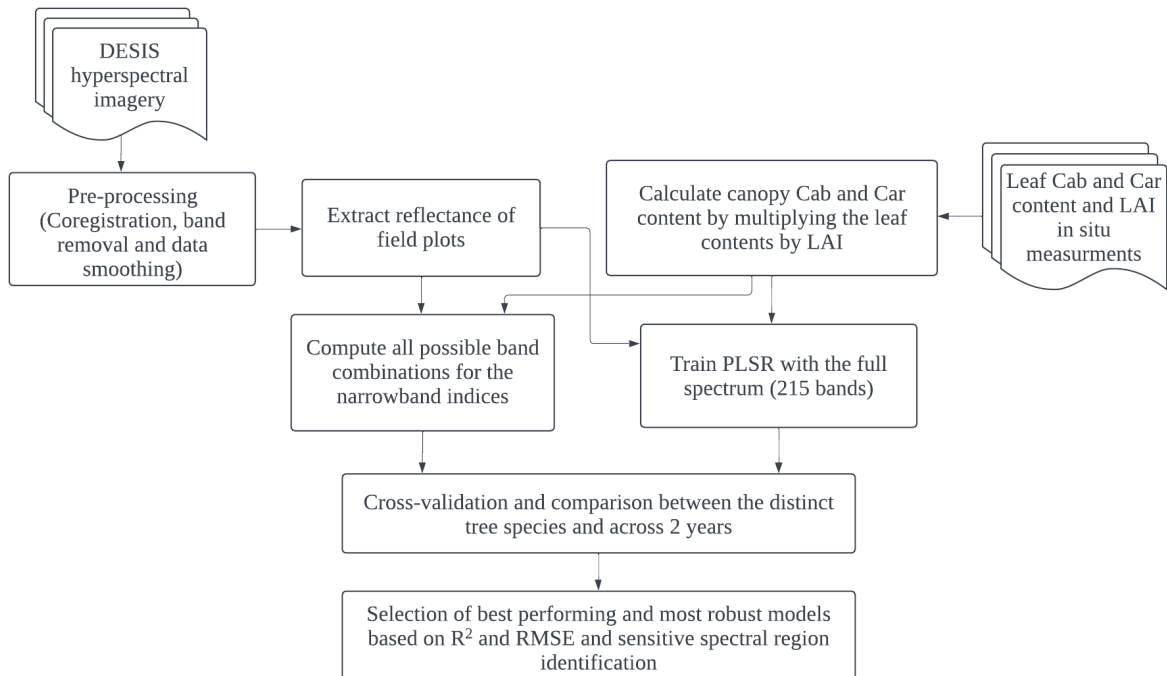


Figure 8. The methodological framework of the study.

3.2. Narrowband vegetation indices

Vegetation indices (VIs) are a combination of reflectance data from two or more spectral bands. Their specific construction provides better reliability than isolated bands (Asner et al., 2003). Additionally, these indices have a diminished response to atmospheric effects and variations in soil brightness (Bannari et al., 1995). Quantifying foliar pigment concentration by using the empirical relations between narrow band optical indices and field samples has successfully been applied to hyperspectral satellite data (e.g., Stagakis et al., 2010; Wu et al., 2008; Wu et al., 2010). The generalized index approach will be employed using the most extensively used and best-performing index formulations. The narrow band vegetation indices were computed using the equation of the indices provided in Table 6.

Table 6. Selected narrowband vegetation indices and their respective formulations.

Index name	Algorithm	References
Simple ratio index (nSRI)	$nSRI = \frac{R_{\lambda 1}}{R_{\lambda 2}}$	Pearson et al., 1972
Modified simple ratio index (nMSR)	$nMSR = \frac{\frac{R_{\lambda 1}}{R_{\lambda 2}} - 1}{\sqrt{\frac{R_{\lambda 1}}{R_{\lambda 2}} + 1}}$	J. M. Chen, 1996
Carotenoid reflectance index (nCRI)	$nCRI = \frac{1}{R_{\lambda 1}} - \frac{1}{R_{\lambda 2}}$	Gitelson et al., 2003
Normalised difference vegetation index (nNDVI)	$nNDVI = \frac{R_{\lambda 1} - R_{\lambda 2}}{R_{\lambda 1} + R_{\lambda 2}}$	Rouse et al., 1974
Optimised soil adjusted vegetation index (nOSAVI)	$nOSAVI = (1 + 0.16) \frac{(R_{\lambda 1} - R_{\lambda 2})}{(R_{\lambda 1} + R_{\lambda 2} + 0.16)}$	Rondeaux et al., 1996
Datt derivative index (nDD)	$nDD = \frac{D_{\lambda 1}}{D_{\lambda 2}}$	Datt, 1999

Pearson et al. (1972) introduced one of the earliest simple ratio indices (SRI), this approach involved using the ratio of near-infrared and red wavelengths for the estimation of leaf area index (LAI). This particular SRI had been identified as a robust technique for assessing canopy chlorophylls by Inoue et al. (2016) who, in their comprehensive review encompassing a variety of vegetation indices and models, proved the efficacy of an SRI that employs a near-infrared wavelength at 815 nm in combination with either 704 nm from the red-edge region or 578 nm from the green region. Addressing the sensitivity of NDVI to various geometrical and optical factors affecting plant canopies, a non-linear function known as the modified simple ratio index (MSR) was introduced. This modification offers a robust alternative to NDVI (Chen, 1996). Derived from the SRI, the carotenoid reflectance index (CRI) was developed to accommodate variations in leaf structure across diverse plant species inhabiting different biomes (Gitelson et al., 2003).

The prevalent indices in use include ratio-based like the normalized difference vegetation index (NDVI) and soil-based such as the optimized soil-adjusted vegetation index (OSAVI). These indices make use of specific red and near-infrared spectral bands, taking advantage of the unique reflectance traits exhibited by vegetation (Darvishzadeh et al., 2008). NDVI is a ratio index based on the normalized difference between the red and near-infrared spectra, since these specific spectral bands are known for their heightened sensitivity to photosynthetically active biomass (Tucker, 1979). OSAVI reduces soil-induced interference on reflectance values in areas with sparse vegetation, this is achieved through the incorporation of a canopy background adjustment factor. OSAVI eliminates the need for prior knowledge of the soil line, making it a simpler alternative to the transformed soil-adjusted vegetation index (TSAVI) (Rondeaux et al., 1996). The Datt derivative index (DD), was designed to mitigate the effects of scatter variations arising from fluctuations in canopy density, leaf structure, and soil background (Datt, 1999b) This method focuses on the curvature and slope of spectral curves instead of individual reflectance values by using the first derivative of the reflectance (Datt, 1999a).

Narrowband vegetation indices were methodically calculated using the canopy spectra by considering all possible combinations of two bands from a dataset comprising 215 bands, spanning the wavelength range of 430 nm to 1000 nm, resulting in a total of 46,010 unique wavelength combinations (215×214) (Darvishzadeh et al. 2008). Subsequently, simple linear regression models were established to explore the associations between these indices for each band combinations and the canopy foliar pigments measured in the field. The most important wavelength combinations for each index were ranked by the coefficient of determination (R^2). Linear regressions were established between the best band combinations and pigments to estimate the pigment concentrations in different species and years.

3.3. Partial least squares regression

One downside of high spectral resolution data is the high dimensionality and collinearity between the explanatory variables (spectral bands) (Hoepfner et al., 2020). Multivariate statistical methods such as multiple linear regression (MLR), principal component regression (PCR), and partial least squares regression (PLSR) can retrieve biophysical and biochemical variables using the whole hyperspectral dataset, overcoming the problem with parametric methods in which the wavelength selection changes depending on the sensor, location, and year (Inoue et al., 2012). Nevertheless, MLR is susceptible to multicollinearity and is greatly impacted by the number of samples and independent variables (Grossman et al., 1996). PCR and PLSR operate similarly and can deal with the high dimensionality and redundancy of hyperspectral datasets (Inoue et al., 2018). However, PCR may be less suitable for predictive purposes as the principal components are solely determined by the variance of the independent variables, while in the PLSR latent variables are defined considering the covariance between the target and independent variables (Norgaard et al., 2000). PLSR has been extensively used as a predictive model for biochemical variable retrieval from hyperspectral data (e.g., Darvishzadeh et al., 2008; Hoepfner et al., 2020; Miraglio et al., 2022). PLSR is a latent variable regression method that iteratively selects linear models that have the maximum correlation with the dependent variable (Inoue et al., 2012).

PLSR can be prone to overfitting, such as other multivariate approaches, when an excessive number of components are selected. This can potentially result in a model that does not include bands causally related to the target variable. Consequently, a regression model might seem to be a good fit and statistically significant but may be specific to the dataset used for model building. Therefore, the selection of the optimal number of components is a critical phase in model fitting. According to Gowen et al. (2011), overfitting is associated with model bias and variance, typically characterized by low bias and high variance in overfitted models. Hence, assessing variance through cross-validation is a conventional practice for determining the optimal number of components to include in the model. In this context, the determination of the optimal number of components for the PLSR model was guided by assessing the root mean square error (RMSE) between measured and predicted foliar pigment values through a leave-one-out cross-validation approach. To avert collinearity and overfitting, the inclusion of components in the model was subject to their ability to reduce cross-validation RMSE by more than 2% (Cho et al., 2007; Darvishzadeh et al., 2008; Kooistra et al., 2004).

The PLSR regression model was developed using the entire set of 215 spectral bands, ranging from 430 nm to 1000 nm. the R 'pls' package was used for model fitting and the MATLAB TOMCAT toolbox (Daszykowski et al., 2007) for verification. Prior to model fitting, the reflectance values were scaled and centred. Scaling involved dividing each variable by its standard deviation to ensure uniform scaling and facilitate model convergence, while centring entailed adjusting the data to have a mean of zero. This preprocessing step enhances the model's robustness and interpretability, allowing for a more effective analysis of the hyperspectral data.

3.4. Validation

Cross-validation is a widely used technique for evaluating the predictive performance of statistical models. In leave-one-out cross-validation (LOOCV), a single data point is reserved for validation, while the remaining $n - 1$ (where n represents the number of observations) are used for iterative model training. LOOCV proves especially beneficial when working with limited datasets where dividing the data into training and test sets is unworkable, as it maximizes the size of the training set. In our study, LOOCV was employed for model validation and ranking. Model ranking relied on cross-validated metrics, including root mean square error (RMSE), normalized root mean square error (NRMSE), and coefficient of determination (R^2) (Ali et al., 2020; Cheng et al., 2017; Ramoelo et al., 2013). The following equations were used:

$$RMSE = \frac{\sqrt{\sum(y - z)^2}}{n} \quad (3)$$

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}} \quad (4)$$

$$R^2 = 1 - \frac{\sum(y - z)^2}{\sum(y - \bar{y})^2} \quad (5)$$

Where y denotes the observed or measured values, z represents the predicted values and \bar{y} is the mean of the observed values.

4. RESULTS

4.1. Narrowband vegetation indices

Among the chosen vegetation indices, the coefficient of determination (R^2) was systematically calculated for every waveband combination to identify the optimal narrowband combination. All indices featuring the best combination presented statistically significant correlations with canopy chlorophylls and carotenoids ($p < 0.01$).

Table 7. RMSE and R^2 of best band combinations using different vegetation indices equation for canopy chlorophylls data across years.

VI	Year	λ_1	λ_2	R^2	RMSE (g/m ²)	NRMSE (%)
nNDVI	2020	760.3	729.4	0.44	0.51	17%
	2021	809.1	767.5	0.26	0.6	8%
nSRI	2020	760.3	742	0.44	0.51	17%
	2021	757.7	752.3	0.33	0.57	10%
nCRI	2020	506.9	568.4	0.23	0.60	9%
	2021	517.2	519.6	0.17	0.64	5%
nMSR	2020	760.3	742	0.44	0.51	17%
	2021	757.7	752.3	0.33	0.57	10%
nOSAVI	2020	760.3	744.5	0.41	0.52	16%
	2021	809.1	767.5	0.34	0.57	11%
nDD	2020	747.2	701.5	0.37	0.54	15%
	2021	757.7	729.4	0.34	0.57	11%

Upon a thorough analysis of the canopy chlorophylls data, it was observed that the most consistent vegetation index across the two years is nDD, with only a 3% difference in the R^2 values (ΔR^2) (Table 7). While others, such as nNDVI, nSRI, and nOSAVI experience a decrease in accuracy in the transition from 2020 to 2021. Based on the average R^2 values, the best-performing models (or vegetation indices) for the canopy chlorophylls data are nSRI, nMSR, nOSAVI, and nDD as they all share a high average R^2 of 0.385 (nSRI and MSR), 0.375 (nOSAVI), and 0.355 (nDD).

Table 8. RMSE and R^2 of best band combinations using different vegetation indices equation for canopy carotenoids data across years.

VI	Year	λ_1	λ_2	R^2	RMSE (g/m²)	NRMSE (%)
nNDVI	2020	760.3	742	0.39	0.09	19%
	2021	755.2	749.7	0.28	0.11	17%
nSRI	2020	760.3	742	0.39	0.09	19%
	2021	757.7	752.3	0.38	0.10	15%
nCRI	2020	506.9	568.4	0.19	0.11	23%
	2021	514.7	524.7	0.22	0.12	18%
nMSR	2020	760.3	742	0.39	0.09	19%
	2021	757.7	752.3	0.38	0.10	15%
nOSAVI	2020	760.3	744.5	0.42	0.09	19%
	2021	809.1	767.5	0.34	0.11	17%
nDD	2020	747.2	716.4	0.35	0.1	21%
	2021	755.2	517.2	0.41	0.1	15%

The nSRI, nMSR, and nCRI indices are the most consistent models for the canopy carotenoids data, as they each exhibit a ΔR^2 of less than 5% between 2020 and 2021. nSRI, nMSR, nOSAVI, and nDD are the top-performing models for the canopy carotenoids data. nSRI and nMSR also demonstrate the highest average R^2 value of 0.385 while for nOSAVI and nDD is 0.380. The most consistent and best-performing models are nSRI and MSR both with a ΔR^2 of 1% and an average R^2 of 0.385.

Table 9. RMSE and R² of best band combinations using different vegetation indices equation for canopy chlorophylls data for different species.

VI	Species	λ_1	λ_2	R ²	RMSE (g/m ²)	NRMSE (%)
nNDVI	Spruce	918.4	972.9	0.18	0.51	8%
	Beech	814.3	770.3	0.27	0.65	9%
	Pine	685.9	683.5	0.2	0.16	31%
	Oak	999.5	724.2	0.25	0.36	15%
nSRI	Spruce	862.9	972.9	0.24	0.49	11%
	Beech	468.3	509.5	0.38	0.6	13%
	Pine	888.1	870.6	0.26	0.15	41%
	Oak	506.9	650	0.4	0.32	24%
nCRI	Spruce	972.9	918.4	0.16	0.51	7%
	Beech	793.1	814.3	0.24	0.66	8%
	Pine	675.8	685.9	0.03	0.18	5%
	Oak	709.4	711.8	0.16	0.39	10%
nMSR	Spruce	862.9	972.9	0.24	0.49	11%
	Beech	816.9	795.9	0.3	0.64	10%
	Pine	888.1	870.6	0.26	0.15	41%
	Oak	506.9	650	0.4	0.32	24%
nOSAVI	Spruce	734.4	732.1	0.20	0.50	9%
	Beech	814.3	770.3	0.25	0.66	8%
	Pine	734.4	732.1	0.20	0.50	31%
	Oak	534.9	599.1	0.29	0.35	17%
nDD	Spruce	644.9	734.4	0.33	0.46	15%
	Beech	501.8	583.7	0.31	0.63	10%
	Pine	655.2	827.2	0.35	0.15	55%
	Oak	683.5	923.9	0.39	0.32	23%

The nSRI and nDD indices emerge as the most accurate models for the canopy chlorophylls data across various tree species, both registering average R² values of 0.32 and 0.345 respectively. The nDD vegetation index demonstrates to be the most consistent model with a ΔR^2 of 6% suggests that its predictive ability is less affected by species variability compared to other indices.

Table 10. RMSE and R² of best band combinations using different vegetation indices equation for canopy carotenoids data for different species.

VI	Species	λ_1	λ_2	R ²	RMSE (g/m ²)	NRMSE (%)
nNDVI	Spruce	918.4	972.9	0.1	0.1	24%
	Beech	757.7	747.2	0.24	0.11	18%
	Pine	685.9	683.5	0.13	0.03	20%
	Oak	509.5	667.9	0.31	0.05	21%
nSRI	Spruce	827.2	972.9	0.18	0.10	24%
	Beech	494	506.9	0.30	0.11	18%
	Pine	890.9	870.6	0.21	0.03	20%
	Oak	506.9	667.9	0.44	0.05	21%
nCRI	Spruce	972.9	918.4	0.06	0.1	24%
	Beech	747.2	760.3	0.21	0.12	20%
	Pine	844.7	890.9	0.05	0.04	27%
	Oak	634.7	522.2	0.15	0.06	25%
nMSR	Spruce	827.2	972.9	0.18	0.1	24%
	Beech	496.5	504.3	0.26	0.11	18%
	Pine	890.9	870.6	0.21	0.03	20%
	Oak	506.9	667.9	0.44	0.05	21%
nOSAVI	Spruce	734.4	732.1	0.11	0.1	24%
	Beech	614.3	770.3	0.24	0.11	18%
	Pine	634.7	642.3	0.11	0.04	27%
	Oak	522.2	690.9	0.3	0.05	21%
nDD	Spruce	460.7	594	0.21	0.09	21%
	Beech	501.8	583.7	0.35	0.11	18%
	Pine	903.7	809.1	0.48	0.03	20%
	Oak	850	688.4	0.37	0.05	21%

No vegetation index consistently performs across different tree species, as all ΔR^2 values exceed the 10% threshold. The nDD shows an overall better performance (lower RMSE and higher R²) for multiple tree species compared to the other indices. This may indicate their versatility and effectiveness for a range of species.

4.2. Most relevant spectral regions

The narrowband Datt derivative index (nDD) consistently outperforms the other indices in terms of accuracy for both years and across all species, as indicated by its lower RMSE and higher R² values. This superior performance points to the effectiveness of the narrowband index in identifying the spectral bands that are most crucial for estimating foliar pigment content. To reveal these key spectral regions, two-dimensional correlation matrices are employed. These matrices graphically display the relationship between the nDD values and the foliar pigment content for each potential combination of bands, using non-cross-validated determination coefficients (R²) values as a metric for correlation strength.

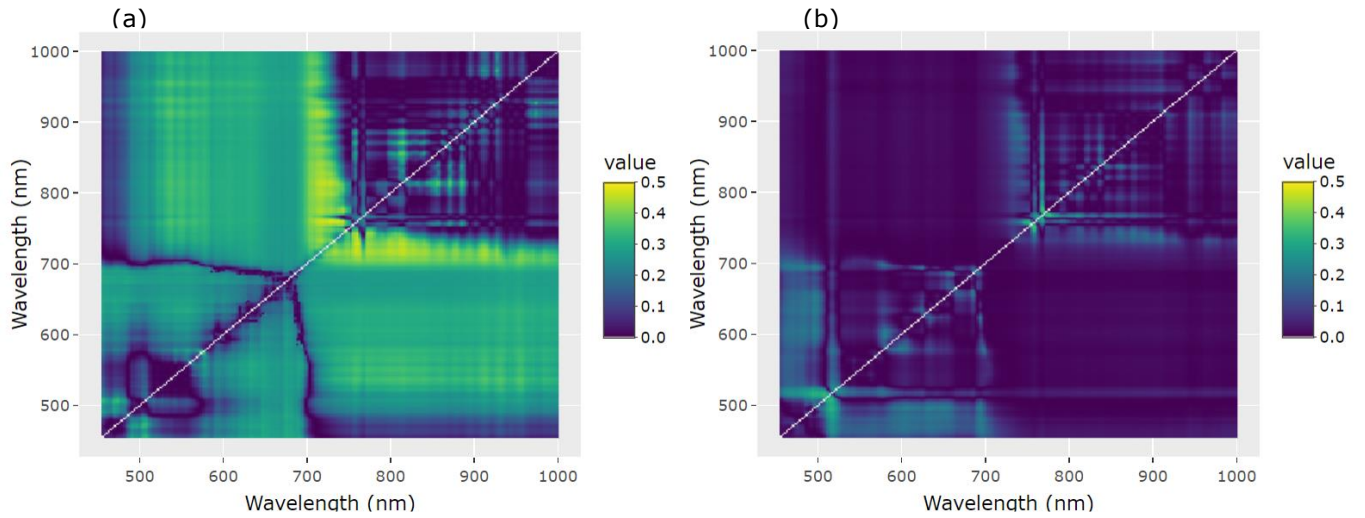


Figure 9. Year-specific 2D correlation plot illustrating the R^2 values between canopy chlorophylls (Cab) content and the narrowband Datt derivative index (nDD) calculated from each spectral band combination for (a) 2020 and (b) 2021.

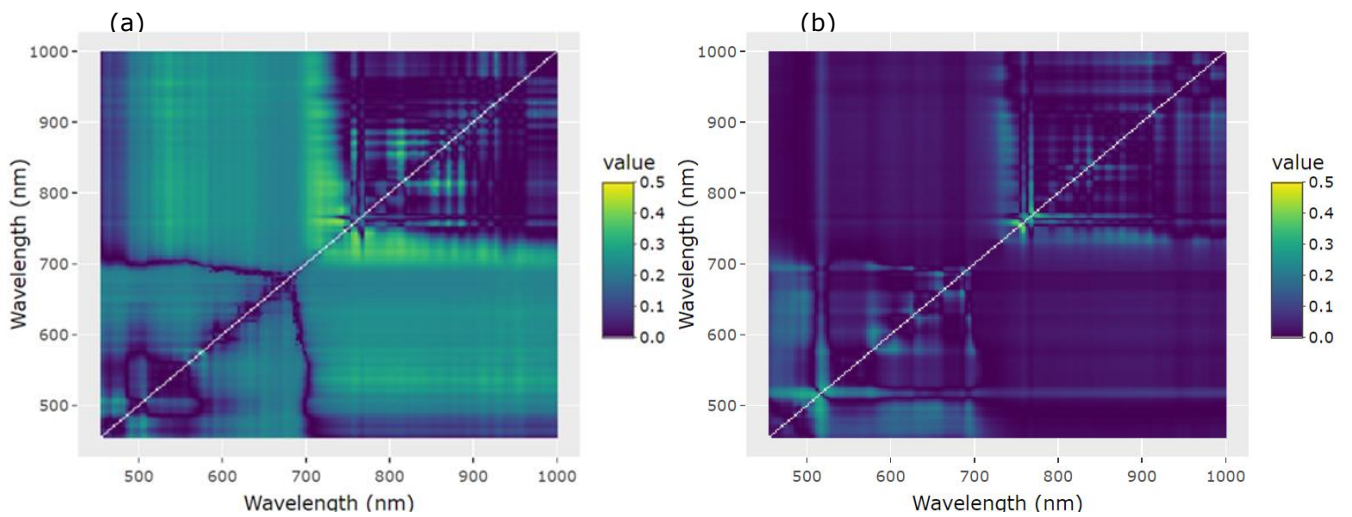


Figure 10. Year-specific 2D correlation plot illustrating R^2 values between canopy carotenoids (Car) content and the narrowband Datt derivative index (nDD) calculated from each spectral band combination for (a) 2020 and (b) 2021.

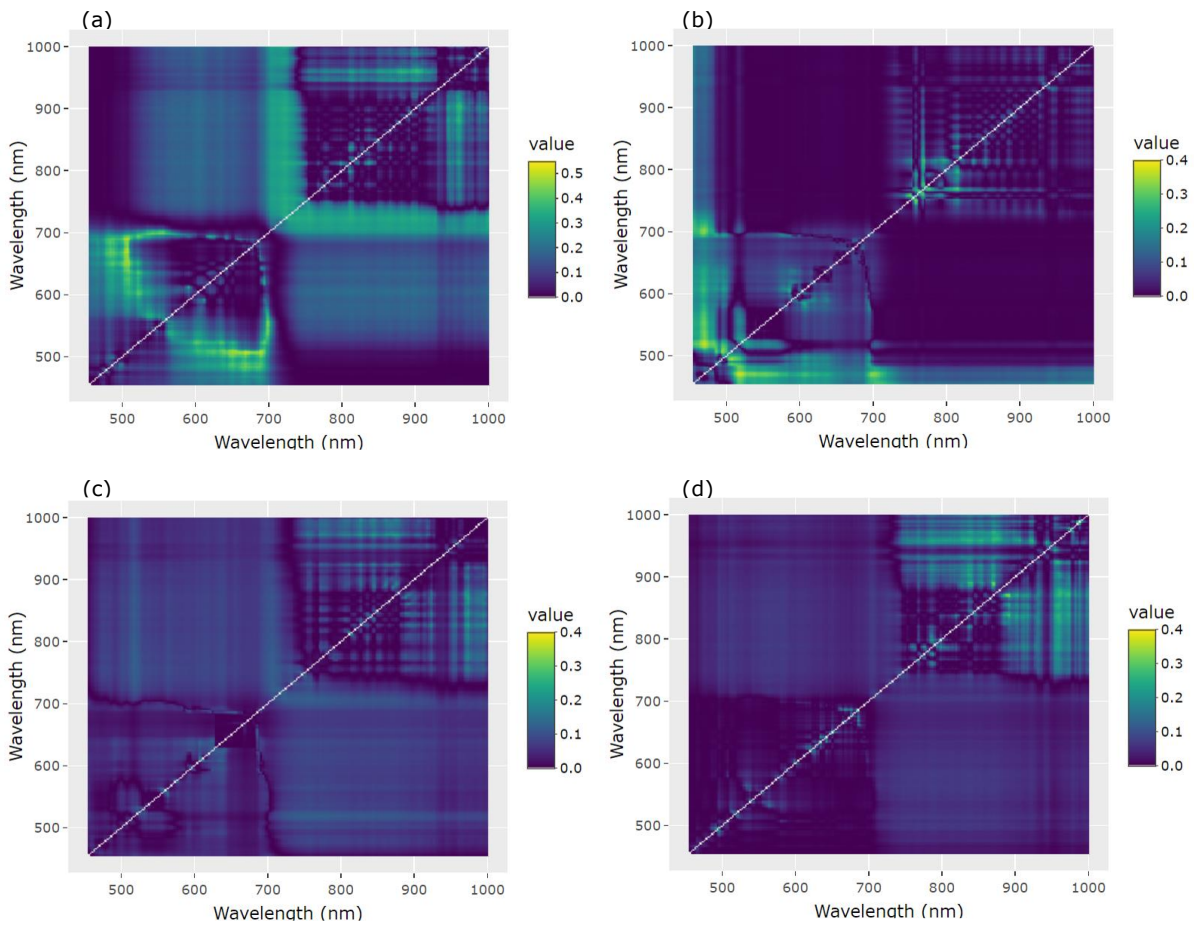


Figure 11. Species-specific 2D correlation plot illustrating the R^2 values between canopy chlorophyll content (Cab) and the narrowband Datt derivative index (nDD) calculated from each spectral band combination for distinct tree species, (a) English oak, (b) European beech, (c) Norway spruce, and (d) Scots pine.

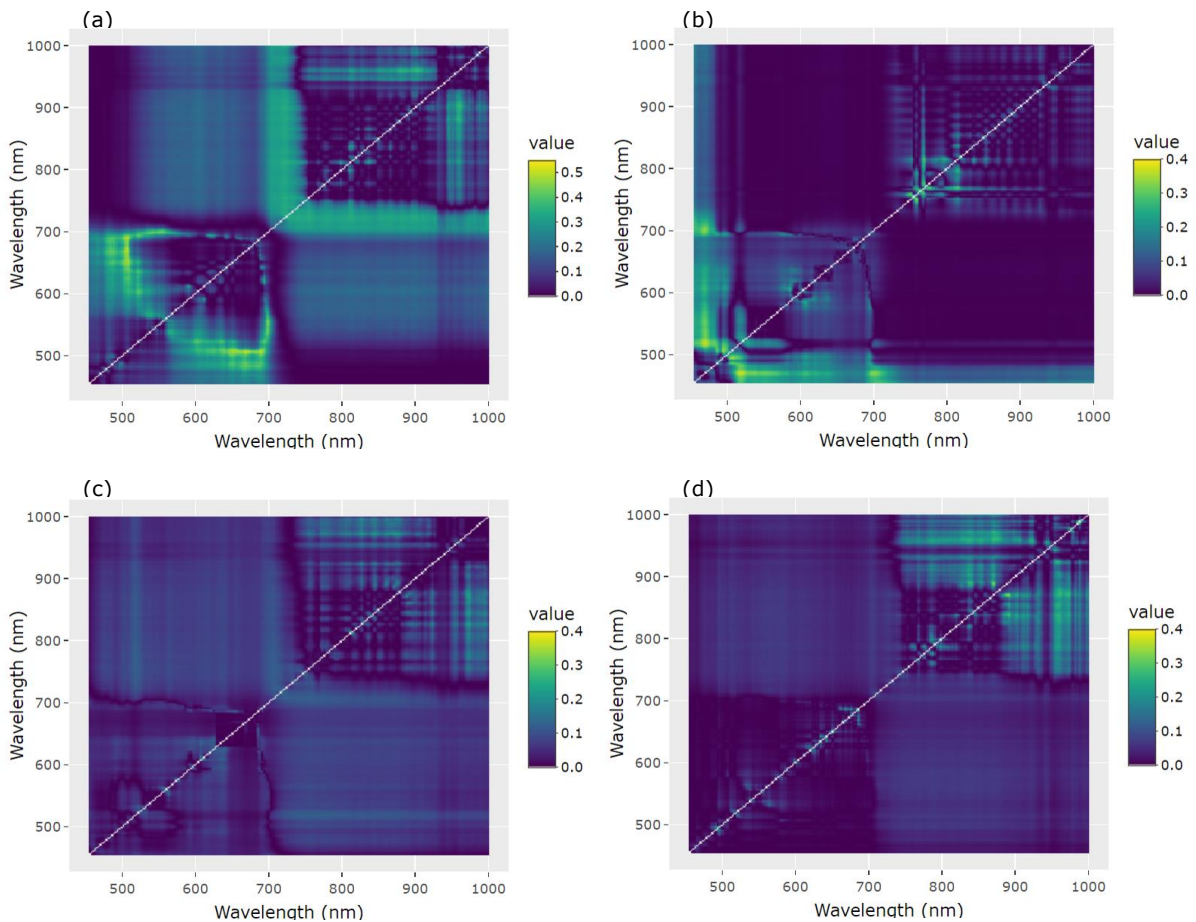


Figure 12. Species-specific 2D correlation plot illustrating R^2 values between canopy carotenoids content (Car) and the narrowband Datt derivative index (nDD) calculated from each spectral band combination for distinct tree species, (a) English oak, (b) European beech, (c) Norway spruce, and (d) Scots pine.

4.3. Partial least squares regression

The results of the partial least squares regression (PLSR) analysis performed for different years and different species are presented below.

Table 11. RMSE and R^2 obtained from PLSR for the estimation of canopy chlorophylls (Cab) and carotenoids (Car) data across two years.

Foliar pigment	Year	Number of samples	Number of components	R^2	RMSE (g/m ²)	NRMSE (%)
Canopy Cab	2020	57	5	0.66	0.40	16%
Canopy Cab	2021	111	5	0.52	0.49	15%
Canopy Car	2020	57	2	0.27	0.10	21%
Canopy Car	2021	111	5	0.53	0.09	14%

The optimal number of components for canopy chlorophyll remained consistent at five across both years. However, the number of components for the canopy carotenoids increased from two in 2020 to five in 2021, suggesting a more complex model or more variability in the 2021 data for canopy carotenoids. For canopy chlorophylls, there was a decrease in R^2 and an increase in RMSE from 2020 to 2021. While for canopy carotenoids, a significant change in R^2 (from 0.27 to 0.53) indicates that the performance of this model also shifted considerably from one year to the next. In conclusion, neither the canopy chlorophylls nor carotenoids model demonstrated stability or consistent performance from 2020 to 2021.

Table 12. RMSE and R^2 obtained from PLSR for the estimation of canopy chlorophylls (Cab) and carotenoids (Car) data across different species.

Pigment	Species	Number of samples	Number of components	R^2	RMSE (g/m ²)	NRMSE (%)
Canopy Cab	Spruce	47	4	0.36	0.44	21%
	Beech	50	5	0.62	0.47	17%
	Pine	23	1	0.06	0.17	27%
	Oak	26	3	0.43	0.31	19%
Canopy Car	Spruce	47	2	0.12	0.09	21%
	Beech	50	5	0.63	0.08	13%
	Pine	23	1	0.06	0.04	27%
	Oak	26	3	0.47	0.04	17%

For both the canopy chlorophylls and carotenoids pigments, the Beech species demonstrated the highest R^2 values, making it the most accurately predicted species in this dataset. Neither the canopy chlorophylls nor carotenoids models consistently perform across all species. The ΔR^2 values surpass the 10% threshold, suggesting variability in the models' performance across different tree species.

The models for both canopy chlorophylls and carotenoids exhibit good predictive capabilities. The canopy chlorophylls models present higher R^2 values, indicating a better fit to the variance in the data. While the canopy carotenoids models have a lower absolute RMSE, the proportionate error for both pigments is comparable when adjusting for the mean pigment values.

4.4. Model performance comparison

In the exploration of the most effective analytical methods, partial least squares regression (PLSR) emerged as the superior technique, demonstrating the highest accuracies both annually ($R^2 = 0.66$ & $RMSE = 0.40$ g/m^2 , for canopy chlorophylls in 2020) and across different species ($R^2 = 0.63$ & $RMSE = 0.08$ g/m^2 , for canopy carotenoids and Beech). Notably, this enhanced accuracy was only observed when the sample size reached a threshold of 50 or more. Contrasting this, the performance of vegetation indices (VIs) appeared to be less sensitive to variations in the number of field samples. Within the scope of VIs, the narrowband Datt derivative index (nDD) and the narrowband simple ratio index (nSRI) were identified as the most effective and reliable indicators. Figure 12 demonstrates the relationships between measured and estimated canopy chlorophyll content using these three leading models, and the comparison of their estimations.

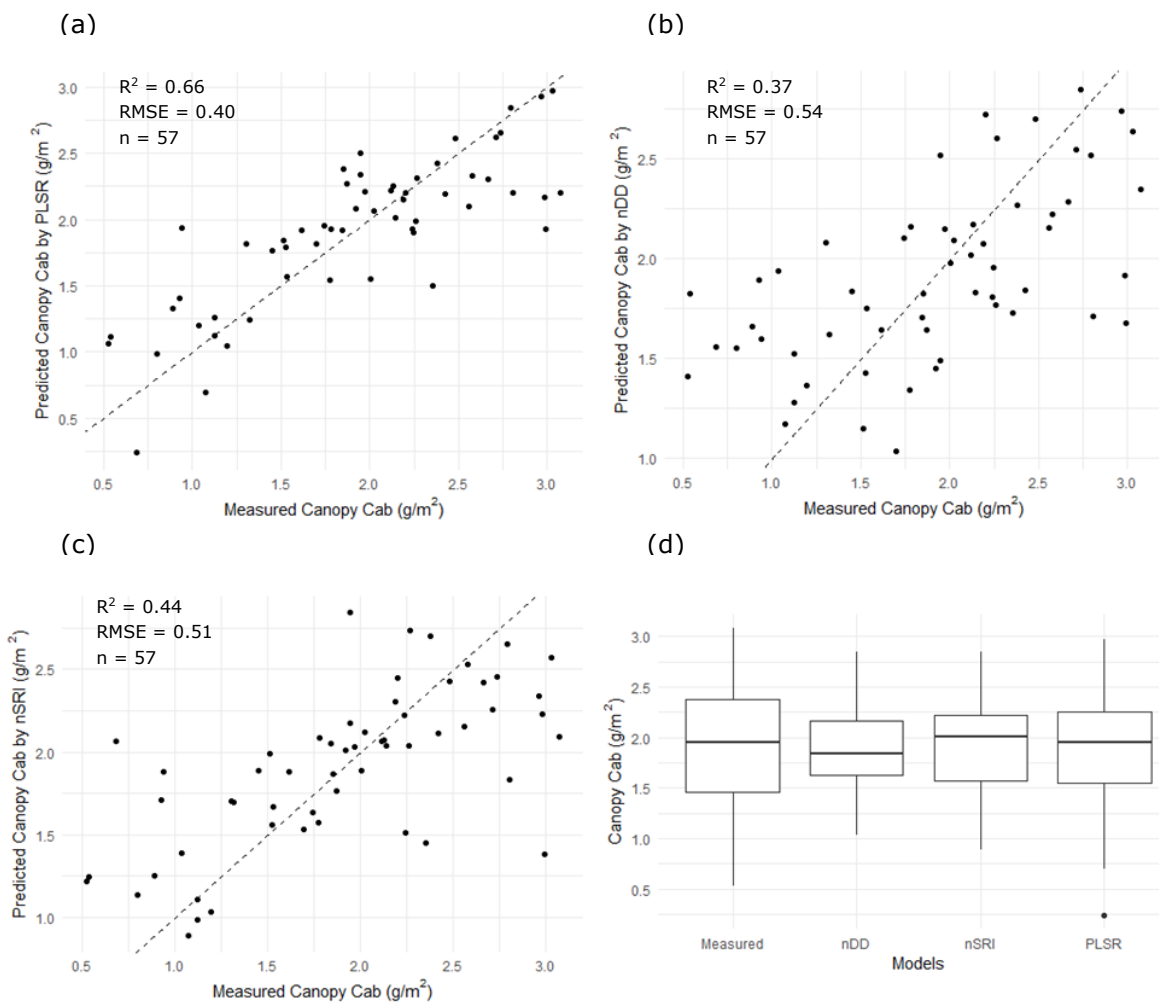


Figure 13. Panels (a) to (c) display scatter plots comparing in-situ measured canopy chlorophylls values with predictions from three different methods for the year 2020, panel (d) features a boxplot of both predicted and measured canopy chlorophylls values. The dashed lines in panels (a) to (c) indicate the ideal 1:1 correlation between predictions and in-situ measurements.

5. DISCUSSION

This study aimed to assess the consistency of canopy chlorophylls and carotenoids content when modelled statistically for temperate forests, while simultaneously testing the accuracy of the model for inter-annual and interspecies variation. Retrieving vegetation traits from canopy reflectance data has multiple challenges, yet the results of this study showed that partial least squares regression (PLSR) demonstrated promising results in estimating the canopy content of chlorophylls and carotenoids for temperate forests. PLSR remained accurate given the natural variation in canopy structure across these protected and unmanaged forests, as well as the radiometric alterations from topography and atmospheric conditions, the varying angles of solar illumination and sensor viewpoints, and the optical properties of soil, which are particularly problematic in areas of sparse vegetation (Darvishzadeh, et al., 2008; Kong et al., 2017; Ollinger, 2011).

5.1. Accuracy of models

In the exploration of the most accurate statistical method, PLSR emerged as the superior technique when compared to the tested vegetation indices, demonstrating the highest accuracies both annually ($R^2 = 0.66$ & $RMSE = 0.40$ g/m², for canopy chlorophylls in 2020) and across different species ($R^2 = 0.63$ & $RMSE = 0.08$ g/m², for canopy carotenoids and Beech), as also demonstrated by other studies (Ali et al., 2020; Darvishzadeh, et al., 2008; Hoepfner et al., 2020). Notably, this enhanced accuracy was only observed when the sample size reached a threshold of 50 or more. In contrast, the performance of vegetation indices (VIs) appeared to be less sensitive to variations in the number of field samples. The accuracy of VIs is notably lower than the PLSR, with coefficients of determination (R^2) lower than 0.50, considered not reliable for predicting foliar pigment content, as categorized by (Chang et al., 2001; Sonobe & Wang, 2017).

Among the narrow band VIs analysed, the nDD consistently excelled in estimating canopy chlorophylls and carotenoids content across varying years and species, also demonstrated in other studies (Hoepfner et al., 2020; Sonobe & Wang, 2017). The narrowband simple ratio index (nSRI) ranked second in accuracy and consistency. All three models predicted similar median canopy chlorophyll as the measured value, exhibiting less variance than the observations, as shown in Figure 12.

5.1.1. Across years

The nDD demonstrates the highest accuracy among all the narrowband VIs, consistently estimating the content of foliar pigments over the observed two-year period. It presented an average R^2 of 0.355 with a ΔR^2 of 3% for canopy chlorophylls and an average R^2 of 0.380 with a ΔR^2 of 6% for canopy carotenoids. A general decline in model accuracy across all VI was noted for canopy chlorophylls from 2020 to 2021, while the results for canopy carotenoids displayed more variability. The influence of the transition between years on the accuracy of these indices appears to be index-specific and pigment-dependent. Some indices, like nSRI and nMSR, maintain relatively stable performance for canopy carotenoids but are not consistent for canopy chlorophylls. Conversely, nNDVI and nOSAVI exhibited a consistent decrease in accuracy from 2020 to 2021 for both pigments.

PLSR showed consistent accuracy values for canopy chlorophylls across the years but a considerable difference in accuracy for canopy carotenoids, with R^2 changing from 0.27 in 2020 to 0.53 in 2021. The increased sample size in 2021 may have contributed to the improved accuracy of the canopy carotenoids predictions, as evidenced by the higher R^2 value. For canopy chlorophylls, the larger sample size did not translate to better model performance, suggesting that factors other than sample size may influence the predictions' accuracy. The number of components used in the PLSR model showed to have an important effect in accuracy. For canopy carotenoids, increasing the number of components from two to five from 2020 to 2021 correlates with improved model performance. PLSR fitting for canopy chlorophylls in 2020, as shown in Figure 12a, revealed a curvilinear relationship, where at higher chlorophyll values the curve approaches an asymptote, suggests reduced accuracy of the model's predictions for higher chlorophylls values. These results highlight the need for careful selection of model parameters and consideration of the unique characteristics of each foliar pigment when applying PLSR for predictive purposes.

5.1.2. Across species

For both canopy chlorophylls and carotenoids, the PLSR and VIs model have a better accuracy with deciduous than with conifer species, aligning with the results found in the literature (Hernandez-Clemente et al., 2014; Miraglio et al., 2019; Navarro-Cerrillo et al., 2014; Sonobe & Wang, 2017). The mean and standard deviation of field measurements for chlorophylls and carotenoids indicate differing levels of pigment content across species. Deciduous trees (European beech and English oak) showed higher variability in both canopy chlorophylls and carotenoids compared to conifers (Norway spruce and Scots pine), as evidenced by larger ranges and standard deviations. This aligns with the notion that the higher variability in these species provides a more robust dataset for these models.

The accuracy of VIs in predicting canopy chlorophylls and carotenoids content varies substantially between species. This variation is likely due to intrinsic differences in the spectral signatures of each species, which are influenced by their unique physiological, biochemical, and canopy structure characteristics. The nDD consistently emerges as one of the most accurate and robust indices for both pigments, across all species. The performance of the indices suggests that species-specific calibrations could be beneficial, especially for coniferous species like Spruce and Pine, where the indices generally perform less effectively.

For PLSR, beech had the highest R^2 values among all species (0.62 for chlorophylls and 0.63 for carotenoids). For the other species, the R^2 was lower than 0.50, considered not reliable for predicting foliar pigment content. The model shows a lower performance for conifer species, with the model's performance for pine being the weakest among the species ($R^2 = 0.06$ for both foliar pigments), having the lowest number of samples ($n = 23$). The number of observations combined with the species type had a considerable influence on model accuracy. While certain species like Beech show promising results with the models, there is evident variability in performance across tree species. Refinements, such as a high number of training samples, or more specific models may be required to enhance predictability across all species, especially for Pine.

5.2. Most important spectral regions for the estimation of canopy chlorophyll and carotenoids content

5.2.1. Annually specific spectral regions

In analysing the canopy chlorophylls data across the two years, a consistent pattern was observed, bands near the spectral region of 760 nm, on the red edge region, were constantly highlighted by multiple narrowband indices, as seen in Figure 8. This region is commonly associated with the nSRI, nMSR, and nOSAVI indices, suggesting its importance in determining canopy chlorophylls content. For the nDD index, particular bands within this zone, specifically the 747.2 nm band in 2020 and the 757.7 nm band in 2021, are identified as relevant for canopy chlorophylls estimation.

Similarly, for canopy carotenoids, the spectral bands near 760 nm and 742 nm are again of interest, consistently involved in the nDD, nSRI and nMSR indices over both years. This spectral range seems to be particularly important for canopy carotenoids estimation, as also demonstrated by Sonobe & Wang (2018). The nDD underlines the 747.2 nm and 701.5 nm bands in 2020 with a notable shift to a more diverse spectral range including the 755.2 nm band and a jump to the green region 517.2 nm band in 2021. Bands around 500 nm and 517 nm were expected to be highlighted (Figure 9), since carotenoids have a maximum absorption peak at around 500 nm (Chappelle et al., 1992).

For both pigment types and across the two-year timeframe, the spectral range of 700-800 nm on the red edge region is prominent (Figures 8 & 9), aligned with what was found in literature since the red edge region has been extensively proven to be very sensitive to foliar pigments (e.g., Hoepfner et al., 2020; Ju et al., 2010; Sonobe & Wang, 2018; Zarco-Tejada et al., 2000). The higher accuracy and consistency of R^2 values from indices like nNDVI, nSRI, nMSR, nOSAVI, and nDD when using bands within this range corroborate its effectiveness for pigment estimation. The consistency of band selection implies that the most important spectral regions for both foliar pigments remain stable across years. These highlighted bands capture unique attributes of the forest canopy.

5.2.2. Species-Specific spectral regions

Detailed analysis of different tree species reveals distinct spectral patterns. For Spruce, the 644.9 nm to 734.4 nm spectral bands were selected by the most accurate index, nDD, for the estimation of the canopy chlorophylls, and canopy carotenoids the 460.7 nm and 594 nm bands were highlighted. Furthermore, the 972.9 nm band and the range around 827 and 863 nm are predominant for Spruce in both the SRI and MSR indices and leaf pigments. In the case of Pine, the nDD points to bands 655.2 nm and 827.2 nm as important for canopy chlorophylls. While for the canopy car, the nDD selected the bands 903.7 nm and 809.1 nm. Across the conifers, there is a notable selection in the range of 645 nm to 655 nm on the red region with both Spruce and Pine displaying preferences in this spectrum for chlorophylls. This observation aligns with the findings of Malenovský et al. (2006), who developed an index to estimate chlorophyll content in Norway spruce. Their research highlighted the significance of the chlorophyll absorption feature within the 650-725 nm wavelength range. Furthermore, the 827.2 nm and 927.9 bands in the NIR are predominant in both the SRI and MSR indices, demonstrating their importance for both pigments. This is further evidenced in Figures 10 and 11, which demonstrate a strong correlation between these NIR regions and the foliar pigments in conifer species. In the NIR region, approximately 750–900 nm, leaf optical properties, including reflectance, are primarily determined by the mesophyll structure rather than direct chlorophyll absorption (Noda et al., 2021).

For Beech, the nDD index distinctly identifies the 501.8 nm and 583.7 nm bands, in the green region, as critical for estimating both canopy chlorophylls and carotenoids. Concurrently, the SRI highlights the 468.3 nm and 509.5 nm bands for chlorophylls, while selecting the 494 nm and 506.9 nm bands for carotenoids. For Oak, the 506.9 nm and 667.9 nm bands, recurrently appear in indices like nSRI and nMSR for both foliar pigments. Oak also seems to be particularly sensitive to the bands on the red region around 683.5 nm and 688.4 nm highlighted by the nDD. The spectral region around 500 nm to 510 nm in the green region and wavelengths between 667.9 – 688.4 nm in the red region are underlined for deciduous species. Research conducted by Falcioni et al. (2023) also highlighted the green/yellow spectral regions, specifically between 500-600 nm, for estimating chlorophylls and carotenoids in Oak and Beech species. This finding aligns with other studies that have similarly identified these wavelength regions as critical for chlorophyll content estimation for deciduous species (Datt, 1998; Gitelson et al., 2003b).

5.3. Possible reasons for discrepancies in results and limitations

Upon examining the in-situ measurements (Table 4), a noticeable change was observed in the mean canopy chlorophylls and carotenoids from 2020 to 2021, showing a decrease in the mean values. The coefficient of variation (CV) for both foliar pigments also increased during this period, potentially indicating a change in environmental conditions between the years. Possible sources are drought or stress induced by insect infestations, such as the multiyear bark beetle infestation that has been ongoing in central European woodlands (Korolyova et al., 2022), affecting pigment concentrations and their uniformity within the canopy (Konôpková et al., 2020). Furthermore, the number of samples in 2021 nearly doubled compared to 2020 for both canopy chlorophylls and carotenoids. This larger dataset in 2021 might offer a broader representation of variability. The difference in sampling locations is also significant, as the number of samples for Veluwe increased from 15 in 2020 to 79 in 2021. Another factor contributing to the discrepancies between years could be the varying dates, times, and sensor positions during image acquisition, as well as differences in image specifications and quality, such as radiometric variations between the images.

An essential aspect to consider is the temporal discrepancy between the in-situ measurements and the DESIS imagery, as detailed in Tables 1 and 2. This time mismatch may lead to notable differences in the data, attributable to changes in environmental conditions or variations in canopy characteristics over time. Furthermore, a significant limitation arises from the plot sizes used for in-situ measurements and the DESIS imagery ground sampling distance (GSD). For optimal accuracy and representation, the plot size for in-situ measurements should exceed the pixel size of the satellite imagery to effectively capture the variability within each pixel.

6. CONCLUSIONS

This research compared established statistical methods, considered species and temporally specific, by training them on diverse species datasets and analysing two years of hyperspectral satellite data. Most studies in the literature mainly focus on improving the already existing models or developing new models, while there is still a gap in the performance of traditional models in varying conditions and on newly available high-spectral resolution satellite imagery.

Answering the first research question of this study, the spectral region around 500 nm to 510 nm in the green region and wavelengths between 667.9 – 688.4 nm in the red region are underlined for deciduous species. Across the conifers, the 645 to 655 nm red spectrum range is key for chlorophyll absorption. Moreover, the 827.2 nm and 927.9 nm bands in the NIR spectrum have proven essential in both SRI and MSR indices for assessing chlorophyll and carotenoid pigments. In analysing the foliar pigment data across the two years, a consistent pattern was observed, bands between 700 nm to 800 nm, on the red edge region were constantly highlighted by multiple narrowband indices.

Addressing the second research question of this study, partial least squares regression (PLSR) and vegetation indices (VIs) have a better accuracy with deciduous than with conifer species for both canopy chlorophylls and carotenoids. PLSR showed the highest predictive accuracy across the statistical models but was year and species-specific. In comparison, some VIs showed consistent results across years and species, the narrowband Datt derivative index (nDD) displays the highest consistency but low accuracy. The accuracy of all VIs is notably lower than the PLSR, with an R^2 lower than 0.50, considered not reliable for predicting the foliar pigment content. Most models show inconsistency across years and species, with the variability between species being greater than between years. Overall, the results showed that the choice of narrowband VIs and the characteristics of the tree species can substantially influence the accuracy of canopy pigment content predictions, as mentioned before in the literature (Ali et al., 2020; Zhen et al., 2021).

The findings of this study highlight the intricacies involved in modelling canopy pigment content, emphasizing the necessity for species-specific statistical methodologies and the integration of multi-temporal data in statistical models. Further research should focus on the robustness and consistency of other remote sensing methods for different species and across time. Additionally, exploring a cross-regional approach is essential to understand how different forest ecosystems across varied geographical landscapes respond to these methodologies, enhancing model adaptability and generalizability. Equally important is the advancements in sensor fusion techniques, which involve integrating data from various sensor types such as LiDAR, multispectral, and hyperspectral sensors. Future studies could enhance retrieval accuracy by synchronizing image acquisition with field data collection, minimizing the time lag between corresponding datasets. The deployment of new hyperspectral satellite sensors like the Hyperspectral Infrared Imager (HyspIRI) and the Copernicus Hyperspectral Imaging Mission for the Environment (CHIME), could significantly improve the retrieval accuracy of canopy pigments content and allow the accurate monitoring of stress and disturbance in forest ecosystems in an era of climate change and biodiversity loss.

LIST OF REFERENCES

- Ali, A. M., Darvishzadeh, R., Skidmore, A., Gara, T. W., O'Connor, B., Roesli, C., Heurich, M., & Paganini, M. (2020). Comparing methods for mapping canopy chlorophyll content in a mixed mountain forest using Sentinel-2 data. *International Journal of Applied Earth Observation and Geoinformation*, *87*, 102037. <https://doi.org/10.1016/J.JAG.2019.102037>
- Asner, G. P., Hicke, J. A., & Lobell, D. B. (2003). Per-Pixel Analysis of Forest Structure. *Remote Sensing of Forest Environments*, 209–254. https://doi.org/10.1007/978-1-4615-0306-4_8
- Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, *13*(1–2), 95–120. <https://doi.org/10.1080/02757259509532298>
- Chang, C.-W., Laird, D. A., Mausbach, M. J., & Hurburgh, C. R. (2001). Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. *Soil Science Society of America Journal*, *65*(2), 480–490. <https://doi.org/10.2136/SSSAJ2001.652480X>
- Chen, J. M. (1996). Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Canadian Journal of Remote Sensing*, *22*(3), 229–242. <https://doi.org/10.1080/07038992.1996.10855178>
- Cho, M. A., Skidmore, A., Corsi, F., van Wieren, S. E., & Sobhan, I. (2007). Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation*, *9*(4), 414–424. <https://doi.org/10.1016/J.JAG.2007.02.001>
- Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F., & Cho, M. (2008). LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS Journal of Photogrammetry and Remote Sensing*, *63*(4), 409–426. <https://doi.org/10.1016/J.ISPRSJPRS.2008.01.001>
- Daszykowski, M., Serneels, S., Kaczmarek, K., Van Espen, P., Croux, C., & Walczak, B. (2007). TOMCAT: A MATLAB toolbox for multivariate calibration techniques. *Chemometrics and Intelligent Laboratory Systems*, *85*(2), 269–277. <https://doi.org/10.1016/J.CHEMOLAB.2006.03.006>
- Datt, B. (1998). Remote Sensing of Chlorophyll a, Chlorophyll b, Chlorophyll a+b, and Total Carotenoid Content in Eucalyptus Leaves. *Remote Sensing of Environment*, *66*(2), 111–121. [https://doi.org/10.1016/S0034-4257\(98\)00046-7](https://doi.org/10.1016/S0034-4257(98)00046-7)
- Datt, B. (1999a). A New Reflectance Index for Remote Sensing of Chlorophyll Content in Higher Plants: Tests using Eucalyptus Leaves. *Journal of Plant Physiology*, *154*(1), 30–36. [https://doi.org/10.1016/S0176-1617\(99\)80314-9](https://doi.org/10.1016/S0176-1617(99)80314-9)
- Datt, B. (1999b). Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. *International Journal of Remote Sensing*, *20*(14), 2741–2759. <https://doi.org/10.1080/014311699211778>
- Falcioni, R., Antunes, W. C., Demattê, J. A. M., & Nanni, M. R. (2023). A Novel Method for Estimating Chlorophyll and Carotenoid Concentrations in Leaves: A Two Hyperspectral Sensor Approach. *Sensors (Basel, Switzerland)*, *23*(8). <https://doi.org/10.3390/S23083843>
- Flannery, B. (2007). *Numerical recipes 3rd edition: The art of scientific computing*.
- Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003a). 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), 271–282. <https://doi.org/10.1078/0176-1617-00887>
- Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003b). 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), 271–282. <https://doi.org/10.1078/0176-1617-00887>
- Gowen, A. A., Downey, G., Esquerre, C., & O'Donnell, C. P. (2011). Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients. *Journal of Chemometrics*, *25*(7), 375–381. <https://doi.org/10.1002/CEM.1349>
- Grossman, Y. L., Ustin, S. L., Jacquemoud, S., Sanderson, E. W., Schmuck, G., & Verdebout, J. (1996). Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data. *Remote Sensing of Environment*, *56*(3), 182–193. [https://doi.org/10.1016/0034-4257\(95\)00235-9](https://doi.org/10.1016/0034-4257(95)00235-9)

- Heurich, M., Beudert, B., Rall, H., & Křenová, Z. (2010). National parks as model regions for interdisciplinary long-term ecological research: The Bavarian forest and šumavá national parks underway to transboundary ecosystem research. *Long-Term Ecological Research: Between Theory and Application*, 327–344. https://doi.org/10.1007/978-90-481-8782-9_23/FIGURES/10
- Hoepfner, J. M., Skidmore, A. K., Darvishzadeh, R., Heurich, M., Chang, H. C., & Gara, T. W. (2020). Mapping Canopy Chlorophyll Content in a Temperate Forest Using Airborne Hyperspectral Data. *Remote Sensing*, 12(21), 3573. <https://doi.org/10.3390/RS12213573>
- Inoue, Y., Darvishzadeh, R., & Skidmore, A. (2018). Hyperspectral Assessment of Ecophysiological Functioning for Diagnostics of Crops and Vegetation. In *Biophysical and Biochemical Characterization and Plant Species Studies*. CRC Press. <https://doi.org/10.1201/9780429431180-2>
- Inoue, Y., Guérif, M., Baret, F., Skidmore, A., Gitelson, A., Schlerf, M., Darvishzadeh, R., & Oliso, A. (2016). Simple and robust methods for remote sensing of canopy chlorophyll content: a comparative analysis of hyperspectral data for different types of vegetation. *Plant, Cell & Environment*, 39(12), 2609–2623. <https://doi.org/10.1111/PCE.12815>
- Inoue, Y., Sakaiya, E., Zhu, Y., & Takahashi, W. (2012). Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. *Remote Sensing of Environment*, 126, 210–221. <https://doi.org/10.1016/J.RSE.2012.08.026>
- Ji, L., Wang, L., & Geng, X. (2019). An automatic bad band pre-removal method for hyperspectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(12), 4985–4994. <https://doi.org/10.1109/JSTARS.2019.2944930>
- Kooistra, L., Salas, E. A. L., Clevers, J. G. P. W., Wehrens, R., Leuven, R. S. E. W., Nienhuis, P. H., & Buydens, L. M. C. (2004). Exploring field vegetation reflectance as an indicator of soil contamination in river floodplains. *Environmental Pollution*, 127(2), 281–290. [https://doi.org/10.1016/S0269-7491\(03\)00266-5](https://doi.org/10.1016/S0269-7491(03)00266-5)
- Li, X., Li, G., & Zhao, B. (2022). Low-Light Hyperspectral Image Enhancement. *IEEE Transactions on Geoscience and Remote Sensing*, 60. <https://doi.org/10.1109/TGRS.2022.3201206>
- Li, Y. H., Tan, X., Zhang, W., Jiao, Q. Bin, Xu, Y. X., Li, H., Zou, Y. B., Yang, L., & Fang, Y. P. (2021). Research and Application of Several Key Techniques in Hyperspectral Image Preprocessing. *Frontiers in Plant Science*, 12, 627865. <https://doi.org/10.3389/FPLS.2021.627865/BIBTEX>
- Lichtenthaler, H. K. (1987). Chlorophylls and carotenoids: Pigments of photosynthetic biomembranes. *Methods in Enzymology*, 148(C), 350–382. [https://doi.org/10.1016/0076-6879\(87\)48036-1](https://doi.org/10.1016/0076-6879(87)48036-1)
- Lichtenthaler, H. K., & Buschmann, C. (2001). Chlorophylls and Carotenoids: Measurement and Characterization by UV-VIS Spectroscopy. *Current Protocols in Food Analytical Chemistry*, 1(1), F4.3.1-F4.3.8. <https://doi.org/10.1002/0471142913.FAF0403S01>
- Malenovský, Z., Ufer, C., Lhotáková, Z., Clevers, J. G. P. W., Schaepman, M. E., Albrechtová, J., & Cudlín, P. (2006). A new hyperspectral index for chlorophyll estimation of a forest canopy: Area under curve normalised to maximal band depth between 650-725 nm. *EARSeL EProceedings*, 5(2), 161–172.
- Miraglio, T., Adeline, K., Huesca, M., Ustin, S., & Briottet, X. (2022). Assessing vegetation traits estimates accuracies from the future SBG and biodiversity hyperspectral missions over two Mediterranean Forests. <https://doi.org/10.1080/01431161.2022.2093143>
- Neefjes, J. (2018). *Landschapsbiografie van de Veluwe*. Rijksdienst voor het Cultureel Erfgoed.
- Norgaard, L., Saudland, A., Wagner, J., Nielsen, J. P., Munck, L., & Engelsen, S. B. (2000). Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. *Applied Spectroscopy*, 54(3), 413–419. <https://doi.org/10.1366/0003702001949500>
- Pearson, R. L., Miller, L. D., Pearson, R. L., & Miller, L. D. (1972). Remote Mapping of Standing Crop Biomass for Estimation of the Productivity of the Shortgrass Prairie. *Rse*, 1355. <https://ui.adsabs.harvard.edu/abs/1972rse..conf.1355P/abstract>
- Rani, N., Mandla, V. R., & Singh, T. (2017). Evaluation of atmospheric corrections on hyperspectral data with special reference to mineral mapping. *Geoscience Frontiers*, 8(4), 797–808. <https://doi.org/10.1016/J.GSF.2016.06.004>

- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55(2), 95–107. [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7)
- Rouse, J., Haas, R. H., Deering, D., Schell, J. A., & Harlan, J. (1974). Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. [Great Plains Corridor]. *NASA/GSFCT Type III Final Report*.
- Schläpfer, D., Richter, R., & 2011, undefined. (2011). *Spectral polishing of high resolution imaging spectroscopy data*. http://www.daniel-schlaepfer.ch/pdf/Schlaepfer_IS2011_polish.pdf
- Sommer, C., Holzwarth, S., Heiden, U., Heurich, M., Müller, J., & Mauser, W. (2015). Feature based tree species classification using hyperspectral and lidar data in the Bavarian Forest National Park. *Elib.Dlr.De*, 2, 49. <https://doi.org/10.12760/02-2015-2-05>
- Sonobe, R., & Wang, Q. (2017). Hyperspectral indices for quantifying leaf chlorophyll concentrations performed differently with different leaf types in deciduous forests. *Ecological Informatics*, 37, 1–9. <https://doi.org/10.1016/J.ECOINF.2016.11.007>
- Stagakis, S., Markos, N., Sykioti, O., & Kyriarissis, A. (2010). Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a *Phlomis fruticosa* Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sensing of Environment*, 114(5), 977–994. <https://doi.org/10.1016/J.RSE.2009.12.006>
- Tsai, F., & Philpot, W. (1998). Derivative Analysis of Hyperspectral Data. *Remote Sensing of Environment*, 66(1), 41–51. [https://doi.org/10.1016/S0034-4257\(98\)00032-7](https://doi.org/10.1016/S0034-4257(98)00032-7)
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- Vaiphasa, C. (2006). Consideration of smoothing techniques for hyperspectral remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(2), 91–99. <https://doi.org/10.1016/J.ISPRSJPRS.2005.11.002>
- Wu, C., Han, X., Niu, Z., & Dong, J. (2010). An evaluation of EO-1 hyperspectral Hyperion data for chlorophyll content and leaf area index estimation. <Http://Dx.Doi.Org/10.1080/01431160903252335>, 31(4), 1079–1086. <https://doi.org/10.1080/01431160903252335>
- Wu, C., Niu, Z., Tang, Q., & Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and Forest Meteorology*, 148(8–9), 1230–1241. <https://doi.org/10.1016/J.AGRFORMET.2008.03.005>
- Zhang, Y., Hui, J., Qin, Q., Sun, Y., Zhang, T., Sun, H., & Li, M. (2021). Transfer-learning-based approach for leaf chlorophyll content estimation of winter wheat from hyperspectral data. *Remote Sensing of Environment*, 267, 112724. <https://doi.org/10.1016/J.RSE.2021.112724>
- Zhu, L., Erving, A., Koistinen, K., Nuikka, M., Junnilainen, H., Heiska, N., & Haggrén, H. (2008). *Georeferencing multi-temporal and multi-scale imagery in photogrammetry*.