# From field to airborne spectroscopy –

*advancing spectral data analytics for accurate retrieval of perennial ryegrass biomass and feed quality*

Gustavo Togeiro de Alckmin

# Propositions

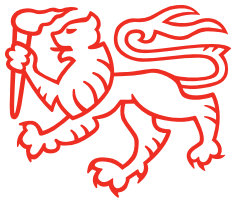
1. Broadscale application of remote sensing for precision pasture management is within technical reach.
   (this thesis)
2. Outdoor spectroscopy has three main and equally important setbacks: measurement error, feature redundancy, and ambiguity in spectral response.
   (this thesis)
3. Science is a play of alternative hypotheses to unveil nature's mysteries.
4. The basic rules of specialization of labor are not applied to a doctoral degree.
5. Science and technology are often mistakenly perceived as synonyms by the public.
6. Complex solutions are often fragile.

Propositions belonging to the thesis, entitled

From field to airborne spectroscopy – advancing spectral data analytics for accurate retrieval of perennial ryegrass biomass and feed quality

Gustavo Togeiro de Alckmin
Wageningen, 19 May 2021 (Date defence ceremony)

# From field to airborne spectroscopy – advancing spectral data analytics for accurate retrieval of perennial ryegrass biomass and feed quality



UNIVERSITY of TASMANIA

WAGENINGEN
UNIVERSITY & RESEARCH

Gustavo Togeiro de Alckmin

**Thesis committee**

**Promotor:**
Dr. Lammert Kooistra
Associate Professor, Laboratory of Geo-information Science and Remote Sensing
Wageningen University & Research

**Co-promotors:**
Prof. Arko Lucieer
Professor of the School of Geography, Planning, and Spatial Sciences
University of Tasmania, Hobart, Australia

Dr. Richard Rawnsley
Researcher of the Livestock Production Centre
Tasmanian Institute of Agriculture, Burnie, Australia

**Other members:**
Prof. Dr. Eldert van Henten, Wageningen University & Research
Prof. Andrew Robson, University of New England, Australia
Prof. Onisimo Mutanga, University of KwaZulu-Natal, South Africa
Prof. Raul Zurita-Milla, University of Twente, Enschede

This research was conducted under the auspices of the School of Geography, Planning and Spatial Sciences, Australia, and the C.T. de Wit Graduate School of Production Ecology & Resource Conservation (PE&RC), The Netherlands, as part of a joint PhD programme.

# From field to airborne spectroscopy – advancing spectral data analytics for accurate retrieval of perennial ryegrass biomass and feed quality

Gustavo Togeiro de Alckmin

**Thesis**

submitted in fulfilment of the requirements for the joint degree of doctor between

**University of Tasmania**

by the authority of the Rector Magnificus, Prof. dr. Rufus Black,

and

**Wageningen University**

by the authority of the Rector Magnificus, Prof. dr. Arthur P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Boards of both universities

to be defended in public

on Wednesday 19 May 2021

at 1:30 p.m. in the Aula of Wageningen University.

Gustavo Togeiro de Alckmin
From field to airborne spectroscopy – advancing spectral data analytics for accurate retrieval of perennial ryegrass biomass and feed quality

Joint PhD thesis, University of Tasmania, Hobart, Australia, and Wageningen University, Wageningen, the Netherlands (2021)
With references, with summary in English

ISBN 978-94-6395-766-3
DOI https://doi.org/10.18174/544521

# Contents

# Chapter 1

# Introduction

## 1.1 Background and Motivation

Grasslands are the most extensive vegetation land cover spanning across 31.5% of the global landmass (Allen et al., 2011; Latham et al., 2014). Adequate grazing-management of these areas allows for the transformation of non-digestible resources into high-value nutritious products, securing sustainable food production systems (Knaus, 2016). In the same way, pasturelands (Allen et al., 2011) are the backbone of the dairy industry in many countries, enhancing the productivity of marginal agricultural lands while supporting rural communities (Hennessy et al., 2020). In Australia and the Netherlands, the dairy sector plays a significant socioeconomic role, generating over AUD$ 4.8 billion (Dairy Australia, 2020) and € 4.9 billion (Beldman et al., 2020) while employing close to 43,500 and 49,000 people, respectively.

In temperate climates, perennial ryegrass (*Lolium perenne* L.) has been widely employed as a major grass species for intensive pasture production (Fulkerson et al., 2007), covering an estimate of six and one million hectares across 5,000 and 14,852 dairy farms in Australia and the Netherlands, respectively (Beldman et al., 2020; Dairy Australia, 2020). In practice, optimal pasture growth is achieved through constant monitoring and decision-making (García et al., 2014). In this way, pasture managers rely on accurate information through extensive spatial and temporal data collection to adjust the rate of inputs (e.g., irrigation and fertilization levels), grazing intervals, and stocking rates. Such decisions have a substantial impact both on pasture quantity and quality (Rawnsley et al., 2014) as well as animal performance (McEvoy et al., 2009).

Often though, despite an extensive body of knowledge outlining best practices, farmers are handicapped in their decision-making process due to the absence of reliable or accessible information about key biophysical and biochemical parameters (Regan et al., 2020). In most cases, pasture management is often conducted through empirical and non-technical practices and can, consequently, lead to poor grazing in the form of overgrazing or low levels of pasture utilization (Hall et al., 2017). Additionally, despite being the main component of the animal's diet, within-farm pasture assessments for feed quality information are usually based on biophysical measurements, morphological proxies, and seasonal trends rather than through analytical procedures (Machado et al., 2005).

In essence, farmers are constrained to laborious methods for biomass estimation while having a limited assessment of pasture quality parameters, usually relying on seasonal patterns (Kallenbach, 2015), while having no direct metric of the consequences of its pasture management practices (Chapman et al., 2012). The absence of this sort of information directly impacts overall farm profitability (Hanrahan et al., 2018) and limits the efficient use of natural resources, leading to environmental impacts (Wachendorf et al., 2004) and non-optimal allocation of scarce resources.

Currently, most techniques for measurement and monitoring of biomass rely on canopy-height measurements (Hall et al., 2017), which have been shown to be prone to inaccuracies (Nakagami, 2016). Such canopy height and biomass relationships were shown to vary throughout season, years and management practices (Thomson et al., 1997). Finally, the continual trend of larger farms, decreasing availability of rural labour and an ageing demographic (Regan et al., 2020) indicates the need for improved and automated methods for pasture assessment (Gargiulo et al., 2018). Requirements such as continuous monitoring of (intra)-paddock quantity and quality are crucial for decision support systems and further developments towards precision pasture management (Jago et al., 2013).

The possibility of fast and automated information retrieval of vegetation is a main and traditional topic within the field of remote sensing (Fussell et al., 1986). Ideally, such techniques can provide accurate information concerning key biophysical and biochemical parameters at an appropriate spatial and temporal resolution for pasture management (Ali et al., 2016). Despite its potential, no effective technological package has yet been devised nor massively adopted by pasture-managers, a main driver for the current marginal impact of precision technologies in pasture-based systems (Shalloo et al., 2018).

Historically, substantial effort has been made to develop spectra-based techniques to monitor grasslands, particularly through satellite data (Rouse et al., 1973) and handheld spectrometers (Pearson et al., 1976). However, the adoption of satellite-based techniques was limited due to the absence of an appropriate spatial and temporal resolution. More recently, the development of portable hyperspectral instruments has fostered many studies concerning the use of spectral data in outdoor environments as a substitute for laboratory Near Infrared Spectroscopy (NIRS) feed quality analysis (Kawamura et al., 2009; Sanches, 2009).

In 1968, as part of a pioneering remote sensing monitoring program, a 10-band multispectral imaging system was mounted on an aircraft and employed to monitor different aspects of grasslands (Tucker et al., 1973; Tucker et al., 1975), as a precursor of satellite observation. There is strong reason to believe that such a program initiated the use of remote sensing techniques for grassland observations, validating the use of satellite observation on a global scale. Within this program another novelty was introduced, Pearson et al. (1976) built a multispectral handheld spectrometer as a tool for fast measurement of plant traits, validating airborne spectral observations and ground-truthing measurements.

At such early period, satellite observation was seen as the only economical platform for regional scale (Verstraete et al., 1996), while field spectroscopy was considered an ideal technique for the retrieval of the true spectral response of a target (Milton, 1987). Such *in-situ* observation could then be employed as a complementary data source for satellite observations or biophysical model validation.

Field spectroscopy is a well-established technology and, although not as accurate as benchtop NIRS, it has been consistently employed for estimations of pasture quantity

(Starks et al., 2006) and quality (Pullanagari et al., 2012). However, despite such technique having shown the great potential of outdoor spectral observations, it has not dealt nor aimed to deal with the issue of automated data-collection. Consequently, most advancements such as narrowband vegetation indices (Mutanga and Skidmore, 2004a), multi-angle spectral data (Gianelle and Guastella, 2007), or multivariate analysis could not be adopted in large scale due to the absence of appropriate platform and pipeline for data collection.

Alternatively, satellite or airborne remote sensing have been constrained by inadequate spatial or temporal resolutions (Schellberg et al., 2008), prohibitive costs and complexity for application-specific tasks (Milton, 1987), such as crop-specific payloads or on-demand revisit intervals. Yet, through the recent popularization of Remotely Piloted Aircraft System (RPAS), a subset of these constraints, particularly those linked with data-collection automation, spatial and temporal resolutions, became manageable (Von Bueren et al., 2015). In summary, the advent of low-cost RPAS allows the possibility of acquiring data at a scale, frequency, convenience and cost, which are suitable for farming operations (Mondino, 2018).

These developments have opened new avenues of opportunity, translating highly complex platforms, such as satellites, into affordable semi-autonomous aerial solutions (Paredes et al., 2017). Nowadays, low-cost Vertical Take off and Landing (VTOL) RPAS (Pircher et al., 2017) can provide the necessary autonomy and range for farm-operations (i.e. below 500 ha), while taking advantage of custom-designed sensing payloads (Price, 1994). In parallel, these custom-designed sensors can benefit from the use of field spectrometers to define optimal bands for the task at hand (Milton, 1987). The selection of a minimal number of bands has impacts on sensor manufacturing costs, data storage and processing power requirements as well as, usually, more robust models. Remarkably, an equivalent miniaturized sensing setup to the 1968 airborne multispectral system can now be designed and deployed in a farm-scenario, at a fraction of the cost (Morales et al., 2020).

Finally, the possibility to automate pasture quantity and quality assessment is in-pace, and to some extent overlap, with new technologies such as virtual fencing, decision support systems and variable rate application of fertilizers. This suite of techniques has the potential to promote precision pasture management and dairy production to higher productivity levels, substantially decrease environmental impact, such as nitrogen leaching, while improving working conditions and attractiveness of rural activities.

In the context of this thesis, the terms Remotely Piloted Aircraft System (RPAS), Unmanned Aerial Vehicle (UAV) and Unmanned Aerial System (UAS) are used interchangeably referring to the aerial platform and sensor, as a unique remote sensing system. However, it is important to dissociate the physical measurement from the platform and data acquisition scale. For such reason, whenever possible, the term low-level flight refers to the spectral observations acquired below 120 meters (400 feet) height, in accordance with the regulations from most civil aviation authorities. Also, the term biomass refers to

aboveground biomass, expressed as kilograms dry-matter per hectare (kg DM/ha). Finally, the term *optimal sensor design* refers to the number and position of bands, rather than the all-encompassing topic of optics and electronics as well as corresponding processes of sensor correction (Kelcey and Lucieer, 2012).

## 1.2 Research Gaps

As a new branch of remote sensing, RPAS remote sensing entails several challenges and research-needs. Firstly, the radiometric correction pipeline (Kelcey and Lucieer, 2012) for an autonomous system operating under changing light conditions (Hakala et al., 2013) and unstable sensor position is still under active research (Suomalainen et al., 2018). Secondly, due to the redundant nature of vegetation reflectance, the task of feature (i.e. band) selection for optimal sensor design is non-trivial (Wiersma and Landgrebe, 1980). The number, width and position of bands have a significant impact on the final prediction accuracy of pasture parameters and sensor cost (Bajcsy and Groves, 2004). Thirdly, the ambiguous response of vegetation spectra provides a poorly determined subset of solutions, imposing limitations in achievable accuracy as different combinations of vegetation parameters can output the same spectral response (Baret and Buis, 2008). These challenges have been sparsely reported throughout the scientific literature, without necessarily discussing optimal solutions such as radiometric correction pipelines, minimal and optimal number of bands in sensors, nor achievable accuracy for each specific retrieval task.

Although a large number of studies in field spectroscopy have dealt with the accurate retrieval of quantitative and qualitative parameters (Kawamura et al., 2009), these studies have not focused on a limitation of practical importance: minimizing the number of necessary spectral bands while maintaining prediction accuracy (Starks et al., 2006). Overcoming such limitation would consequently lead to a straightforward translation into low-cost, simple sensors, readily applicable to RPAS remote sensing and farming scenarios.

Most of these studies (Kawamura et al., 2009; Sanches, 2009; Starks et al., 2006) directly translated laboratory-based NIRS methods to field spectroscopy, as well as employing chemometric techniques, such as pre-processing techniques and spectral transformations, which may present limited utility in multispectral systems or in outdoor-scenarios due to the inherent low signal-to-noise ratio. Consequently, these approaches, although valid in their context, do not address measurement errors nor the ambiguous nature of outdoor spectral measurements (Baret and Guyot, 1991).

Alternatively, it seems rational to understand the limitations of spectra-based methods in uncontrolled scenarios (Verstraete et al., 1996) and determine achievable accuracies, while minimizing sensor complexity and cost (i.e., number of bands) rather than naively

aiming to minimize error metrics through non-robust models. Furthermore, employing the informative part of the spectra, rather than the full output of hyperspectral sensors should be beneficial both to the modeling task and the understanding of the phenomena (Kjeldahl and Bro, 2010).

Besides an optimal sensor design through the selection of spectral bands, a relevant unattended gap refers to the definition of appropriate modeling strategies (Verrelst et al., 2015) for the different pasture parameters, as well as validation strategies (Meyer et al., 2019) which would allow insights into model performance in different contexts (e.g., location, years or season). The ability to better understand and quantify these gaps, providing evidence of model performance across different scenarios is crucial for broad-scale adoption of remote sensing techniques in agricultural activities (Meyer et al., 2018), and assessment of the benefits of these new upcoming technologies.

Although a large number of studies have dealt with achievable accuracies for a particular biophysical or biochemical parameter, validation strategies may have been *"naive"* in their prevention of issues such as data leakage, leading to overfit and poor results whenever employed in different scenarios (Meyer et al., 2019). Additionally, within the last decades, an overall trend for complex black-box models has increased model fit at the cost of lower levels of interpretability and inference (Roscher et al., 2020), while under such naive validation strategies. These challenges and shortcomings, which are commonly approached through a data-science point-of-view, are dealt and discussed across the thesis.

## 1.3 Problem Statement

Although many empirical studies have aimed to determine the achievable accuracy of spectra-based models when mapping perennial ryegrass biophysical and biochemical characteristics, none have specifically aimed to determine the best and smallest subset of bands to achieve such accuracies. This task is not trivial and has not been clearly described in the scientific literature.

Furthermore, most models are not assessed under distinct validation strategies, providing poor evidence of reproducibility of results and robustness across different contexts. Consequently, there is a significant gap between understanding whether spectral models are generalizable outside the boundary conditions (i.e., across different sites and dates) under which data was collected.

Finally, it is important to determine whether the achievable accuracy found through handheld data is equivalent to spectra-based models in low-level flight and whether these accuracies are sufficient for pasture-management, justifying the adoption of such.

Provided that these gaps are better or completely understood, sufficient evidence can be provided to determine the benefits of RPAS remote sensing against the current standard

methods for pasture management. Also, non-destructive and accurate monitoring of the biophysical and biochemical properties of perennial ryegrass provides an effective technique for phenotyping purposes. The integration of continuous pasture monitoring with other upcoming digital technologies can provide farmers with key performance indicators, significantly reducing inputs while maintaining or enhancing productivity.

In summary, if robust remote sensing techniques can be made available and ready for pasture-managers adoption, a substantial impact can be made on productivity. Consequently, this can meet the upcoming global demand for food production (i.e., food security), improve the use of marginal lands, and encourage pasture conversion to croplands (Zalles et al., 2019) whenever suitable.

## 1.4 Research Aim and Objectives

The overarching aim of this thesis is to expand and enable the use of spectral data for perennial ryegrass management, bridging the techniques of field spectroscopy and optimal imaging systems. To this end, this thesis employs a range of regression algorithms, band selection procedures, spectral transformations and validation strategies to enable timely, repeatable and cost-effective imaging systems for pasture management.

To achieve the main aim, a set of four objectives were developed and employed throughout the thesis. Table 1.1, shows in which chapter each research objective is addressed:

- **Objective 1** – To evaluate the trade-offs between subsets of spectral bands and achievable accuracies for retrieval of biophysical and biochemical parameters.
- **Objective 2** – To assess the impact of different regression algorithms for an optimal model fit between spectral features and pasture parameters.
- **Objective 3** – To determine model performances through the use of rigorous validation protocols for different temporal and geographical circumstances.
- **Objective 4** – To explore the transferability of spectra-based models from canopy to low-level flight.

A series of progressive studies are presented, discussing the superiority of spectra-based models in relation to standard practices for biomass assessment as well as the optimal representation of biochemical parameters, the retrieval and redundancy of hyperspectral data for broadband absorption features and the transferability of handheld methods to low-level sensing systems.

Previous research-gaps which have been elucidated refer to the achievable accuracy for biomass and crude protein retrieval based solely on top-of-canopy canopy reflectance, model-performance in different locations and years, as well as judicious selection of spectral ranges and regression algorithms.

**Table 1.1:** Research Objectives and their relationship with thesis chapters

| Chapters | Objectives | Titles |
|---|---|---|
| **2** | 1 & 2 | Comparing methods to estimate perennial ryegrass biomass: canopy height and spectral vegetation indices. |
| **3** | 1 & 3 | Retrieval of crude protein in perennial ryegrass using spectral data at the canopy level. |
| **4** | 1 & 3 | Retrieval of hyperspectral information from multispectral data for perennial ryegrass biomass estimation. |
| **5** | 2, 3 & 4 | Perennial ryegrass biomass retrieval through multispectral UAV data. |

## 1.5 Thesis Structure

In order to fulfill the objectives of this thesis, four experimental sites were established for data collection. The main experimental site was located in Elliot (TAS, Australia), while three other experimental sites were located in the Netherlands (Goutum, Vredepeel, and Zegveld). While the experimental site in Elliot was used for all the studies in this thesis, the locations in the Netherlands were specifically included for chapters 3 and 4. Data collection campaigns largely employed both handheld hyperspectral spectrometers, RPAS multispectral cameras and complementary instruments for canopy-height estimation.

The chapters of this thesis are based on four peer-reviewed publications, each addressing one or more research objectives as described above.

**Chapter two** critically addresses the use of vegetation indices and canopy height for biomass estimation while quantifying accuracy improvements based on different regression algorithms. As a proxy for *status-quo* techniques, a comparison between canopy height models and the normalized vegetation index (NDVI) is performed. In addition, to further explore the potential of vegetation indices, a brute-force procedure was employed to generate 11,026 normalized ratio indices (NRI) while selecting the best NRI band combination. In parallel, a pool of 97 vegetation indices (synthesized from the literature) was filtered and underwent a feature selection procedure to determine an optimal small subset of indices. Several regression algorithms were evaluated as well as whether significant differences in prediction performance could be attributed to different algorithms.

In **chapter three**, a genetic algorithm was employed in a two-objective search procedure: to minimize the number of spectral bands while simultaneously maximizing model accuracy for crude protein estimation (CP). This protocol was employed over different spectral ranges, namely VIS-NIR (400–1100 nm), SWIR (1100–2500 nm) and the Full-Spectrum range (400—2500 nm), while comparing achievable accuracies of two different metrics:

crude protein as dry matter fraction (% CP) or in a weight-per-area basis (kg CP/ha). Model performance was assessed whenever the validation set covered all locations and whenever a new and unknown location was used as the validation set.

**Chapter four** presents a new approach for retrieval of a continuous spectral signature (550–790 nm) from discrete multispectral measurements (i.e., four bands, as per a commercially available multispectral camera) based on a piecewise function described by two parametric sub-functions. The retrieval of spectral signatures allowed for the generation of continuum-removed features and associated vegetation indices for the prediction of biomass, previously reported in the literature as optimal indices for biomass estimation. These synthetic vegetation indices were compared against vegetation indices derived from the original band values, and model performance for each subset of indices was contrasted. Finally, model performance was assessed for known and unknown locations.

**Chapter five** employed RPAS multispectral imagery, handheld spectral-data and five distinct decision-rule regression techniques to validate the approach of biomass assessment employing a small subset of indices, while critically addressing the challenges in radiometric calibration, model interpretability, model deployment in an operational scenario, and model-performance through different validation strategies. The five regression algorithms build upon the concept of regression-trees, using techniques of bootstrapping aggregation (i.e., bagging) and boosting, consequently increasing model complexity (e.g., number of trees, depth of trees) while decreasing overall interpretability. The RPAS multispectral measurements was compared against handheld top-of-the canopy spectral measurements and significant inconsistencies were found between reflectance values of both sensors. The absence of well-defined and robust protocols for spectral data collection of commercial multispectral cameras is discussed, and predictive performance between handheld and low-level flight spectral is evaluated.

In the **chapter six** of the thesis, General Discussion and Outlook, special emphasis is directed to maximum achievable accuracies through the use of spectral data in outdoor environments, under the presence of confounding and masking effects derived from canopy geometry and illumination conditions while stressing the need for rigorous protocols and quality assurance mechanisms for spectral data collection. Finally, this chapter identifies the main bottlenecks for the advancement of spectral imaging techniques in a farm-operational scenario, indicating possible advancement through optimal multispectral sensing equipment, automated data-collection, faster and interpretable modeling techniques.

# Chapter 2

# Comparing Methods to Estimate Perennial Ryegrass Biomass: Canopy Height and Spectral Vegetation Indices

This chapter is based on: G. Togeiro de Alckmin, L. Kooistra, R. Rawnsley, and A. Lucieer (2021). "Comparing methods to estimate perennial ryegrass biomass: canopy height and spectral vegetation indices". *Precision Agriculture* 22.1, 205–225. DOI: 10.1007/s11119-020-09737-z. URL: http://link.springer.com/10.1007/s11119-020-09737-z

# Abstract

Pasture management is highly dependent on accurate biomass estimation. Usually, such activity is neglected as current methods are time-consuming and frequently perceived as inaccurate. Conversely, spectral data is a promising technique to automate and improve the accuracy and precision of estimates. Historically, spectral vegetation indices have been widely adopted and large numbers have been proposed. The selection of the optimal index or satisfactory subset of indices to accurately estimate biomass is not trivial and can influence the design of new sensors. This study aimed to compare a canopy-based technique (rising plate meter) with spectral vegetation indices. It examined 97 vegetation indices and 11026 combinations of normalized ratio indices paired with different regression techniques on 900 pasture biomass data points of perennial ryegrass (*Lolium perenne*) collected throughout a one-year period. The analyses demonstrated that the canopy-based technique is superior to the standard normalized difference vegetation index ($\Delta$ 115.1 kg DM/ha RMSE), equivalent to the best performing normalized ratio index and less accurate than four selected vegetation indices deployed with different regression techniques (maximum $\Delta$ 231.1 kg DM/ha). When employing the four selected vegetation indices, random forests was the best performing regression technique, followed by support vector machines, multivariate adaptive regression splines and linear regression. Estimate precision was improved through model stacking. In summary, this study demonstrated a series of achievable improvements in both accuracy and precision of pasture biomass estimation, while comparing different numbers of inputs and regression techniques and providing a benchmark against standard techniques of precision agriculture and pasture management.

## 2.1 Introduction

Efficient pasture production and utilization is often the most critical component in a pasture-based dairy operation (García et al., 2014). Despite the extensive literature available about the constituent factors of pasture production and management, most pasture-based grazing systems are not optimally managed due to the costly and time-consuming nature of standard methods for pasture measurement and monitoring. In practice, coordinating pasture growth rates and grazing events is crucial to optimize yield (Chapman et al., 2012) while avoiding nutritional value losses (Turner et al., 2006).
In a farm scenario, such coordination can be achieved through weekly measurements of pasture biomass (García et al., 2014). Currently, common methods, such as rising plate meter (RPM), rely on linear relationships between canopy height (Allen et al., 2011) and biomass. However, these relationships are limited in their accuracy and biased due to plant-development stages, canopy architecture (erectophile or plagiophile) or canopy density (Nakagami, 2016). Finally, such a method requires a trained observer constantly monitoring and sampling paddocks, which can lead to inconsistencies due to observer-bias (Thomson et al., 1997) and requires extensive time investment (Hall et al., 2019).
Alternatively, *remote sensing* (RS) techniques can be employed to provide semi- or fully-automated monitoring, provided that the ideal temporal-spatial scales are observed. This constraint has, historically, prevented widespread adoption, as satellite revisit intervals and pixel size are not adequate for on-farm management. Furthermore, persistent cloud cover is a pervasive constraint (Kawamura et al., 2008) during seasons of rapid growth rates, hence, decreasing the utility of satellite optical imagery (Ali et al., 2016).
*Unmanned aerial vehicles* (UAV) and *multispectral* (MS) cameras (those defined as multi-camera 2D imagers in Aasen et al. (2018)), provide a flexible and economical system for spectral observation at a time and spatial-scale which are ideal for agricultural practices. As an advantage, such multispectral cameras are readily available, producing radiometric calibrated imagery which can be easily integrated into agricultural management tools. A drawback, however, is that multi-camera 2D imagers are mostly restricted to a small number of bands, usually five as per Aasen et al. (2018), within the *visible to near-infrared* range (VIS-NIR) of the spectrum due to their silicon-based sensors. As a consequence, end users have focused on *vegetation indices* (VI) as enhanced predictors for biomass estimation.
Vegetation indices have been, to an extent, successfully employed in *precision agriculture* (PA) and RS to estimate different biophysical and biochemical attributes of vegetation. Accordingly, a vast number of VIs have been proposed (Xue and Su, 2017), surpassing the possibility of optimally or correctly utilizing each one of them without extensive domain knowledge. This issue is found in UAV remote sensing, where most commercial sensors rely on a broadband VI such as the *normalized difference vegetation index* (NDVI) to estimate a range of different, non-correlated attributes. Notwithstanding its versatility, estimations from a multi-purpose VI (i.e., NDVI) are, logically, not optimal for attributes different

from which the index was originally designed. Issues such as NDVI saturation, occurring after a certain biomass threshold, are well-described in the literature (Thenkabail et al., 2002).
Despite the saturation drawback with the traditional NDVI, Mutanga and Skidmore (2004a) have shown that, through systematic generation of *normalized ratio indices* (NRI) for all band pairs of a hyperspectral dataset (acquired at handheld level), certain narrowband NRIs can overcome biomass saturation. In parallel, Burkart et al. (2014) demonstrated the equivalence between spectral observations at two different data acquisition scales: handheld and low-level flight. In a more recent study, Wang et al. (2019) presented a brief review on the difference between reflectance measurements from both acquisition-scales, indicating that such differences are negligible. Such findings fulfill a methodological gap for data collection, analysis and performance validation for UAV sensors (i.e., low-level flight) extrapolated from handheld spectral information.
Presently, these approaches (i.e., filtering and selection of optimal NRIs and known VIs) have not been tested and translated to custom-made commercial UAV sensors. Hence, canopy based approaches, such as the RPM, are still favored in pasture-based systems rather than spectroscopy methods despite the unique potential, at a low-cost, offered by the adoption of UAVs, multispectral sensors and a satisfactory small set of VIs.
Given this research-gap, current commercial multispectral cameras and data analysis (Michez et al., 2019) have reported poor performances for pasture biomass estimation, both when employing spectral-based ($R^2 = 0.35$) or canopy based (photogrammetric estimated height) techniques ($R^2 = 0.23$). This is in spite of the potential reported in previous research (Mutanga and Skidmore, 2004a) and absence of scaling-up issues (Burkart et al., 2014; Wang et al., 2019).
The research objective of this study was to evaluate how many and which VIs (i.e., *features*) should be employed for perennial ryegrass (*Lolium perenne*) biomass estimation while comparing with two traditional techniques: the rising plate meter and broadband NDVI. Consequently, accuracy improvements can be quantified based on the nature of measurements (canopy height or spectra) and number of indices. Furthermore, different regression techniques (parametric and non-parametric algorithms) were analysed, so that increments in accuracy and precision of feature(s)-algorithm pairs can be evaluated.

## 2.2 Methods

### 2.2.1 Experimental Design

The trial was undertaken at the Tasmanian Dairy Research Facility in Elliot (TAS, Australia - 41°04'57.3" S, 145°46'21.8" E). The experimental layout was an array of 30 rainfed perennial ryegrass plots (dimensions of 2.0 x 7.5 m), with 0.35 m border at each side of the plot's longitudinal axis), arranged as two rows by 15 columns (Figure 2.1).

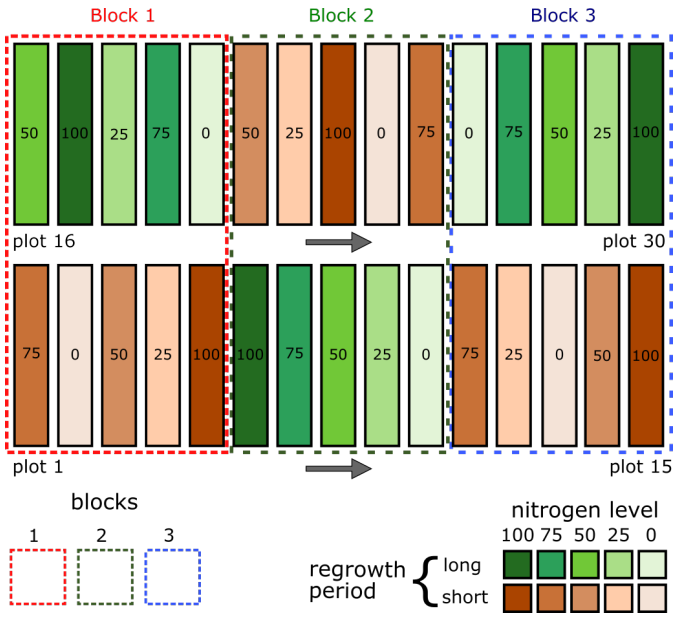



**Figure 2.1:** Experimental layout employed in each of the five data collection campaigns. Colour-codes refer to regrowth periods (colours) and nitrogen levels (hue). Each block (encased by a dashed rectangle) comprises all combinations of regrowth periods and N fertilizing regimes. Plots are numbered from left-to-right and bottom to top. Within, each plot, six different sample-sites were measured both by the spectroradiometer and the rising plate meter

Plots were grouped in three main blocks (10 plots per block). Each block was split in two different growth intervals: long and short or approximately 30 and 15 days, respectively. Each plot on the split-block was randomly allocated a different nitrogen (N) fertilizing regime (0, 25, 50, 75 or 100 kg N/ha). The fertilizer was manually applied (i.e., top-dressing) on each plot at the start of each regrowth cycle, having urea as N source. Prior to spring (end of August) and prior to installing the experiment, phosphorus (P), potassium (K) and sulfur (S) were broadcast throughout the trial area according to soil analysis to ensure that the lack of macronutrients would not impede pasture growth.

Data collection campaigns consisted of three subsequent stages: (1) spectral measurement, (2) canopy height measurement and (3) biomass determination. In each of these stages, attention was given to minimize confounding factors and ensure independence amongst measures, given that measurements of stages (1) and (2) overlapped spatially (Figure 2.2). In total, five campaigns were carried out from December 2016 to November 2017 (as per the dates of spectral measurements).

1. **Spectral measurements**. Spectral data was collected by a field spectroradiometer (ASD Handheld 2, Colorado, USA) on five dates under clear-sky conditions and around solar noon: December 18$^{th}$ 2016, February 06$^{th}$, April 29$^{th}$, October 22$^{nd}$ and November 28$^{th}$ 2017. This instrument acquires data from 325 to 1075 nm, with a total of 750 bands and field of view (FOV) of 25°. Total time spent to obtain all

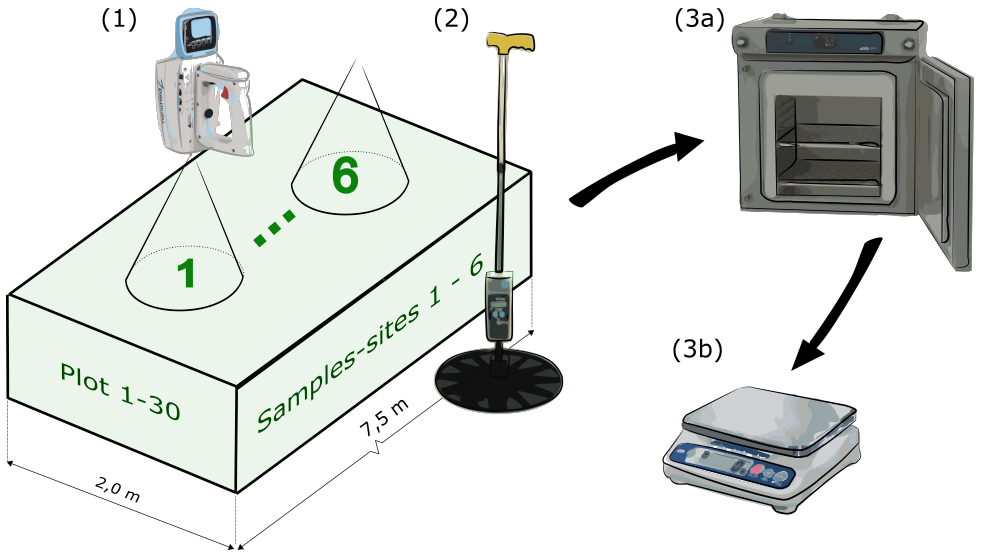



**Figure 2.2:** Data collection workflow. Stages: 1 - Spectral measurements (spectroradiometer); 2 - Canopy height measurements (rising plate meter); 3 - Biomass determination, a) drying (forced-air oven), b) weighing (digital scale). Plots (1-30) dimensions: 2.0 x 7.5m. Sample-sites (1-6) were measured both with the spectroradiometer and the rising plate meter

spectral measurements (180 data points) was 1.5 to 2 h per field campaign with minimum warm-up of 30 minutes. The instrument setup follows the manufacturer's recommendation: 30 scans for spectrum averaging, 60 scans for dark current and white reference. The sequence of measured plots was randomized to minimize any systematic effect of solar position across the plots during data collection. In addition, after finishing measuring the samples of each plot, a spectral measurement of the white reference (Spectralon®) was recorded. The intention of this procedure was twofold: (a) to monitor the stability of the instrument and (b) detect any possible change in atmospheric conditions. The instrument was recalibrated (against the white reference) after seven minutes of continuous usage or whenever the white-reference measurement deviated from 100% reflectance, whichever occurred first. Within each plot, six randomly allocated sample-sites were selected (Figure 2.2 - 1). Spectral measurements were taken from approximately one-meter height, thus, yielding a circular footprint equal to 0.15 $m^2$ (or 0.44 m diameter). Each sample-site was measured five times. Final sample spectral value (referred to as raw spectral data) was the average value of these five measurements.

2. **Canopy height measurements**. An analogue RPM with 5 mm resolution, described in Earle and McGowan (1979), was employed to measure (compressed) canopy height, once per sample-site (Figure 2.2 - 2).

3. **Biomass determination**. Pasture biomass was mechanically defoliated above a residual height of 50 mm from the 0.15 $m^2$ footprint used in stage 1 and 2 (Figure 2.2

- 1 and 2). Harvested material was dried for a minimum of 48 hours at 60 °C in a forced-air oven (Figure 2.2 - 3a) immediately following each harvest for pasture *dry-matter* (DM) determination. Samples were weighed (Figure 2.2 - 3b) using a digital scale (MassCal, 30 kg ±0.5 g).

### 2.2.2 Data Analysis

For reproducibility purposes, data analysis operations are introduced by the corresponding *package::function* format (italics typeface and accompanied by the double colon operator, i.e., the *scope resolution operator*).

Data analysis was performed in RStudio/R (versions 1.14 and 3.5.1, respectively). Necessary packages for the analysis, besides the base and dependencies packages, are *hsdar* (Lehnert et al., 2018), *caret* (Kuhn, 2008) and *caretEnsemble* (Mayer and Knowles, 2015).

### 2.2.3 Feature Generation

The raw spectra data were smoothed using the Savitzky–Golay filter (*hsdar::smoothSpeclib*, window-size = 9 nm, polynomial-degree = 2). This operation aims to decrease the influence of random instrument noise without distorting the original reflectance values. As a consequence, this operation improves the signal-to-noise ratio.

***Vegetation Indices***. The smoothed spectra were transformed into a set of 97 VIs (*hsdar::vegindex*) available in the literature. The full list of VIs can be found in Lehnert et al. (2018) and are listed in Table 2.2. As a baseline approach and as a complementary step for data analysis, all VIs were fitted in a univariate (single-index) ordinary linear regression against biomass values.

***Normalized Ratio Indices***. The smoothed resampled spectra were used in an exhaustive-search process, testing all available bands combinations in a normalized difference equation, as described in Mutanga and Skidmore (2004a) and shown in Eq. 2.1. The smoothed spectra were resampled from 750 to 149 bands given that the necessary computation time to test all band combinations would be substantial, and most likely superfluous due to high spectral band correlation. Spectra were resampled (*hsdar::spectralResampling*) applying a Gaussian response function and 10 nm bandwidth, reducing computational load to 4% of all possible combinations. In total, 11,026 combinations were assessed.

$$NRI(\lambda_1, \lambda_2) = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \tag{2.1}$$

All the NRIs were then fitted in a univariate (single-index) ordinary linear regression (*stats::lm*) against biomass values. From these, the best fitting linear model (*hsdar::nri_best_performance*, highest $R^2$) was identified and the best performing NRI ($\lambda_1$,$\lambda_2$ - *optimized NRI*) included in the pool of filtered VIs.

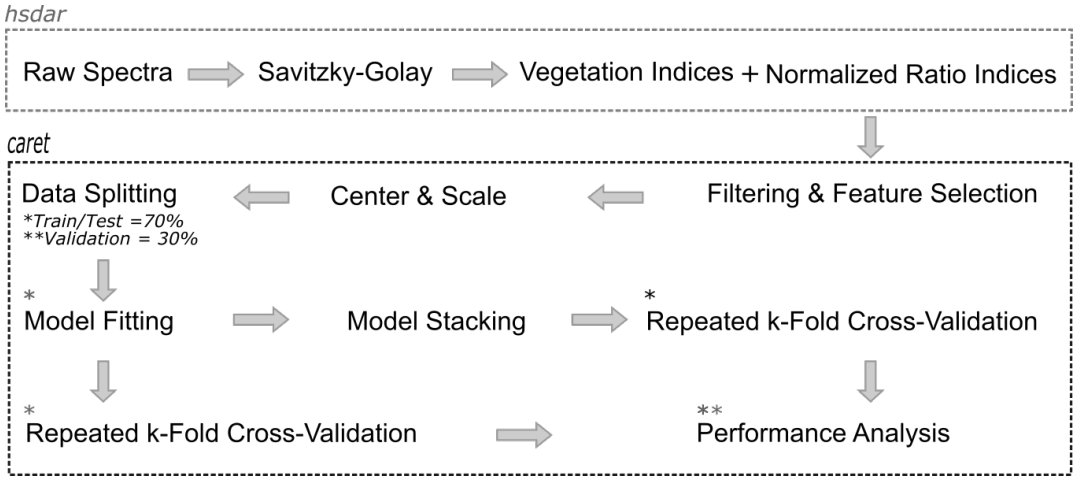



**Figure 2.3:** Data analysis workflow. Statistical packages employed are indicated on the top-left of each dashed rectangle. Feature generation process is enclosed by the top dashed rectangle while feature filtering, model fitting and performance analysis are encased in the lower dashed rectangle. Single and double asterisks indicate where train/test and validation dataset were, respectively, employed.

### 2.2.4 Feature Filtering and Selection

The framework for data analysis consisted of three steps: (A) *filtering* of highly correlated and non-significant VIs (i.e., features); (B) *recursive feature elimination* and feature selection (optimal and satisfactory subsets) and (C) *model fitting, stacking and validation*. A detailed description and rationale of the feature filtering and selection workflow is discussed in Perez-Riverol et al. (2017) and a specific case-study for perennial ryegrass is available in Alckmin et al. (2019). The training/testing set ($n = 630$) comprises 70% of the entire dataset and was used in a repeated $k$-fold cross-validation (folds = 10 and repeats = 5).

A) ***Filtering***: Pearson correlation among all VIs was calculated. A maximum cut-off of |0.95|, based on a sensitivity analysis developed in Alckmin et al. (2019), was applied to identify highly correlated VIs. Such VIs were evaluated in a pair-wise fashion: the one with the largest mean correlation (i.e., correlation with all other features) was removed. A minimum Pearson correlation removal cut-off equal to |0.2| between the remaining filtered VIs and DM values was applied.

B) ***Recursive Feature Elimination***: after the filtering process, remaining VIs were centered, scaled and a recursive feature elimination process (*caret::rfe*) was performed using the training set. Two subsets were identified: optimal (minimal RMSE) and the satisfactory subset, which was the smallest group of features that presented results (in training-testing stages) which were below a 10% threshold from the minimum RMSE model (optimal subset). This workflow and guidelines were presented in Kuhn and Johnson (2013) and Perez-Riverol et al. (2017).

Ranking of VIs (variable importance, *caret::varImp*) for each different VI subset size, at

each cross-validation fold, was calculated through a random forests routine using Gini Importance, introduced in Breiman et al. (2017).
C) **Model Fitting, Stacking and Validation**: after the satisfactory feature set was determined (i.e., *Selected VIs*), different regression algorithms were fitted to the data. The following models were chosen: "*Bagged MARS*" (Friedman, 1991), "*Random Forests*" (Breiman, 1984) and "*Support Vector Machines with Polynomial Kernels*" (Cortes and Vapnik, 1995) and "*Ordinary Linear Regression*" (referred to hereafter as MARS, RF, SVM and LM, respectively). The rationale for this selection was to test models that do not share the same core technique or are simply variations of the same technique. Tuning of hyper-parameters was performed automatically at the training-test stage (*caret::train*) through an embedded grid-search algorithm. Model ensemble/stacking consisted of training (*caretEnsemble*) an algorithm to combine the predictions of other previously trained algorithms (Caruana et al., 2004). Such an approach was generated through a generalized linear model of the three non-parametric models (referred to as *STACK*).
Finally, models were validated against an unseen dataset, corresponding to 30% of total observations ($n = 270$), where it is expected that model performance is similar (or superior) to the training-test stage. For benchmarking purposes, *RPM*, *NDVI*, *optimized NRI* and the satisfactory subset (i.e., *Selected VIs*) were all fitted using a linear model and underwent the *Performance Analysis* protocol, allowing a comparison between features.

### 2.2.5 Performance Analysis

In this study, the error of each feature(s)-algorithm pair, or the difference between predicted and observed/true values (sample weight), was assessed using three different metrics: *root-mean-square error* (RMSE), *coefficient of determination* ($R^2$) and *mean absolute error* (MAE). Ultimately, this analysis aimed to (i) assess the accuracy (Joint Committee for Guides in Metrology, 2008) for each feature(s)-algorithm pair; (ii) to test if the accuracy differences between the feature(s)-algorithm pairs were statistically significant ($p$-value$>0.05$); (iii) if the error-distribution (i.e., precision) of these feature(s)-algorithm pairs were statistically different (Figure 2.4).
For items (i) and (ii), the algorithm (*caret::diff.resamples*) and workflow provided in Kuhn and Johnson (2013) was employed, which are descriptions of benchmark studies mostly based in one sample t-test. To check if the error-distribution was different between feature(s)-algorithm pairs, item (iii), the Kolmorogov-Smirnov test (Massey, 1951) was employed. The two-sample Kolmogorov-Smirnov (*KS*) test is a non-parametric analysis that compares the cumulative distributions of two datasets. In this context, it was employed using error metrics (RMSE) derived from each feature(s)-algorithm pair. The error-metrics ($n = 50$) were generated through a repeated $k$-fold cross-validation (5 folds, 10 repeats) on the training set. Feature(s)-algorithm pairs are presented in Figure 2.4 under the *"Model Fitting"* section. A schematic diagram of this workflow is presented in Figure 2.4.

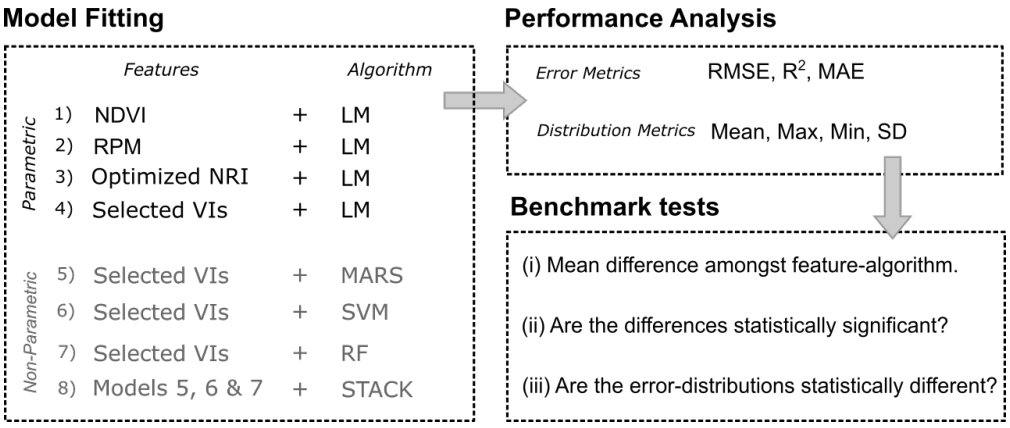



**Figure 2.4:** Feature(s)-algorithm pair model fitting (left-side), performance and tests workflow (right-side). Selected VIs were used as features on non-parametric models. Non-parametric models were used as features on the STACK model

## 2.3 Results

### 2.3.1 Experimental Setup - biomass and canopy height:

Pasture biomasses ranged from 164 kg DM/ha (minimum) to 4663 kg DM/ha (maximum) and a mean of 1633 kg DM/ha. The descriptive statistics data of biomass and canopy height values, per campaign, can be found in Table 2.1.

**Table 2.1:** Descriptive statistics for biophysical characteristics per campaign.

| | Canopy Height (mm) | | | | Biomass (kg DM/ha) | | | |
|---|---|---|---|---|---|---|---|---|
| Date | Average | SD | Max | Min | Average | SD | Max | Min |
| Dec-16 | 124.2 | 31.1 | 225 | 75 | 1837.3 | 677.0 | 4222.2 | 322.3 |
| Feb-17 | 65.2 | 20.3 | 115 | 30 | 1115.5 | 829.8 | 3801.3 | 164.4 |
| Apr-17 | 56.8 | 19.8 | 115 | 20 | 1117.4 | 471.4 | 2380.7 | 269.6 |
| Oct-17 | 129.8 | 27.8 | 205 | 50 | 2450.1 | 943.2 | 4662.8 | 559.0 |
| Nov-17 | 99.2 | 37.7 | 210 | 45 | 1646.2 | 877.4 | 3761.8 | 401.2 |
| **Total** | **95.1** | **41.0** | **225** | **20** | **1633.3** | **923.0** | **4662.8** | **164.4** |

### 2.3.2 Feature Generation

***Vegetation Indices:*** The result of a baseline approach (i.e., single index linear regression), and consequent model fit metrics, can be examined in Table 2.2. The optimized NRI

(*"Opt NRI"*), retrieved within this analysis ranks as the best performing VI (Table 2.2) as found by Mutanga and Skidmore (2004a).

**Table 2.2:** Performance by vegetation index in an ordinary linear regression. Bold indices are selected within the satisfactory subset. Underlined indices are those remaining after the filtering process.

| Vegetation Index | $R^2$ | RMSE | Vegetation Index | $R^2$ | RMSE |
|---|---|---|---|---|---|
| **Opt NRI** | **0.64** | **530.3** | PARS | 0.46 | 651.5 |
| RDVI | 0.63 | 538.7 | SR2 | 0.45 | 654.3 |
| SAVI | 0.63 | 541.9 | Datt | 0.45 | 654.6 |
| MSAVI | 0.61 | 553.8 | mNDVI | 0.45 | 657.7 |
| OSAVI | 0.60 | 557.0 | TCARI2/OSAVI2 | 0.45 | 658.7 |
| OSAVI2 | 0.60 | 557.3 | SR | 0.44 | 660.4 |
| DD | 0.60 | 559.5 | Vogelmann3 | 0.44 | 661.5 |
| MCARI2/OSAVI2 | 0.60 | 562.4 | PRI norm | 0.43 | 668.5 |
| Vogelmann2 | 0.59 | 564.9 | PSRI | 0.42 | 676.9 |
| Vogelmann4 | 0.59 | 570.4 | PRI | 0.41 | 681.5 |
| MCARI2 | 0.59 | 570.7 | SIPI | 0.40 | 685.8 |
| Datt2 | 0.57 | 578.1 | NPCI | 0.40 | 688.4 |
| MTVI | 0.57 | 579.4 | SRPI | 0.39 | 694.7 |
| Sum Dr2 | 0.57 | 583.2 | **SR7** | **0.38** | **695.1** |
| SPVI | 0.56 | 588.7 | Datt5 | 0.38 | 698.0 |
| TVI | 0.56 | 588.9 | GDVI 2 | 0.38 | 698.7 |
| Sum Dr1 | 0.56 | 589.9 | NDVI3 | 0.37 | 704.2 |
| Vogelmann | 0.55 | 591.9 | SR5 | 0.36 | 707.9 |
| Carter4 | 0.55 | 592.5 | Datt4 | 0.35 | 714.2 |
| Green NDVI | 0.55 | 596.0 | DWSI4 | 0.33 | 727.9 |
| SR6 | 0.54 | 597.8 | GI | 0.32 | 729.1 |
| NDVI2 | 0.54 | 600.5 | SR4 | 0.32 | 731.4 |
| mND705 | 0.54 | 600.6 | Boochs | 0.31 | 733.7 |
| mSR705 | 0.54 | 602.9 | MCARI | 0.31 | 736.6 |
| Boochs2 | 0.54 | 603.9 | SR8 | 0.29 | 745.3 |
| REP Li | 0.53 | 604.3 | Carter5 | 0.29 | 745.7 |
| mSR2 | 0.53 | 606.8 | GDVI 3 | 0.28 | 753.0 |
| D1 | 0.53 | 607.4 | TCARI2 | 0.28 | 754.1 |
| EGFR | 0.53 | 607.9 | Datt6 | 0.25 | 766.4 |
| MTCI | 0.53 | 609.7 | MPRI | 0.24 | 770.3 |
| GMI1 | 0.52 | 612.0 | **CI** | **0.24** | **773.5** |
| SR3 | 0.52 | 612.0 | Carter | 0.23 | 777.9 |
| CI2 | 0.52 | 613.2 | GDVI 4 | 0.20 | 790.8 |
| EGFN | 0.51 | 617.7 | TGI | 0.17 | 806.6 |
| PWI | 0.51 | 620.4 | TCARI | 0.17 | 808.5 |
| GMI2 | 0.51 | 621.1 | CRI1 | 0.11 | 834.2 |
| SR1 | 0.51 | 621.1 | mSR | 0.10 | 842.7 |
| Carter3 | 0.50 | 623.5 | MCARI/OSAVI | 0.09 | 843.6 |
| Gitelson2 | 0.50 | 625.5 | Gitelson | 0.09 | 845.6 |
| Carter2 | 0.50 | 626.9 | ClAInt | 0.07 | 854.3 |
| **DDn** | **0.49** | **629.8** | CRI2 | 0.06 | 860.4 |
| Maccioni | 0.49 | 630.6 | DPI | 0.05 | 864.9 |
| PSND | 0.49 | 635.1 | TCARI/OSAVI | 0.04 | 867.6 |
| mREIP | 0.48 | 637.8 | Carter6 | 0.03 | 874.1 |
| Datt3 | 0.48 | 638.8 | EVI | 0.02 | 876.4 |
| NDVI | 0.48 | 640.3 | CRI3 | 0.01 | 880.6 |
| REP LE | 0.48 | 641.6 | PRI*CI2 | 0.01 | 880.6 |
| PSSR | 0.47 | 645.9 | CRI4 | 0.01 | 881.0 |
| D2 | 0.46 | 649.0 | CARI | 0.00 | 886.0 |

**Table 2.2:** Performance by vegetation index in an ordinary linear regression. Bold indices are selected within the satisfactory subset. Underlined indices are those remaining after the filtering process. *(continued)*

| Vegetation Index | $R^2$ | RMSE | Vegetation Index | $R^2$ | RMSE |
|---|---|---|---|---|---|

*Obs:* Values corresponding to an ordinary linear regression (*stats::lm*).

***Optimized Normalized Ratio Indices:*** Figure 2.5 presents the $R^2$ value for each NRI ($\lambda_1$,$\lambda_2$ combination) and pasture biomass values. The region with the highest correlation to pasture biomass (delimited by a dashed rectangle) occurred within the far end of the red-edge region (680–730 nm) until the NIR shoulder (up to 900 nm). The best performing NRI ($\lambda_1 = 745$, $\lambda_2 = 755$ nm) is then added to the feature-space in which the feature selection will be performed.

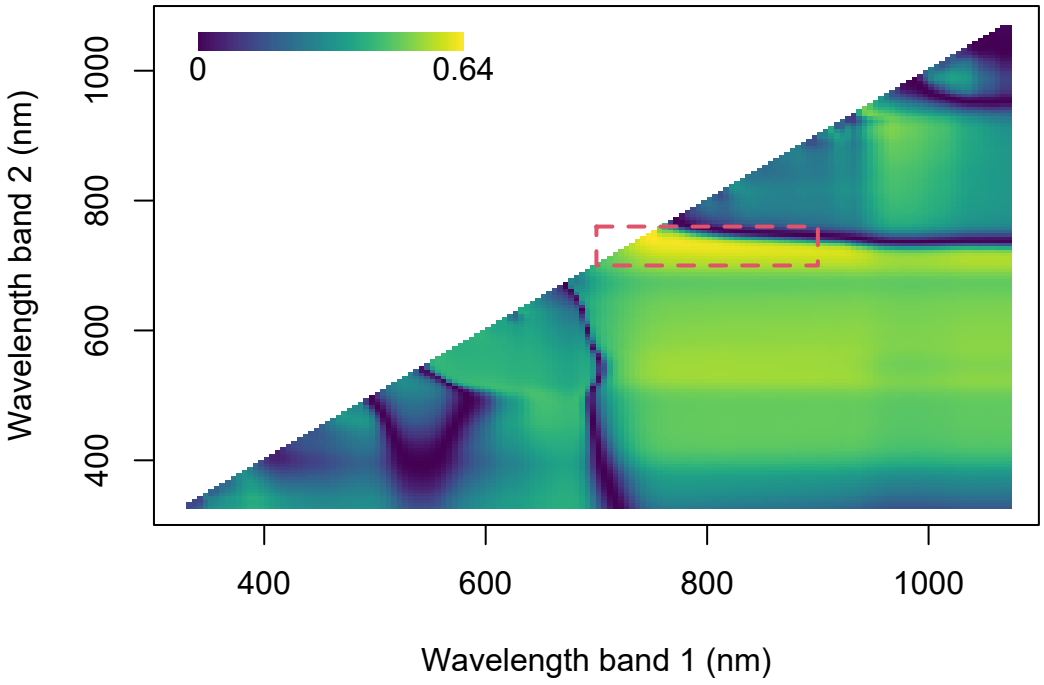



**Figure 2.5:** Normalized Ratio Index Correlogram ($\lambda_1$, $\lambda_2$ and biomass). Displaying the determination coefficients between biomass and narrow band NRI values calculated from all possible combinations spread across $\lambda_1$ (350—1075 nm) and $\lambda_2$ (350—1075 nm). Dashed rectangle limits $\lambda_1$ 700–900 and $\lambda_2$ 700–760 nm

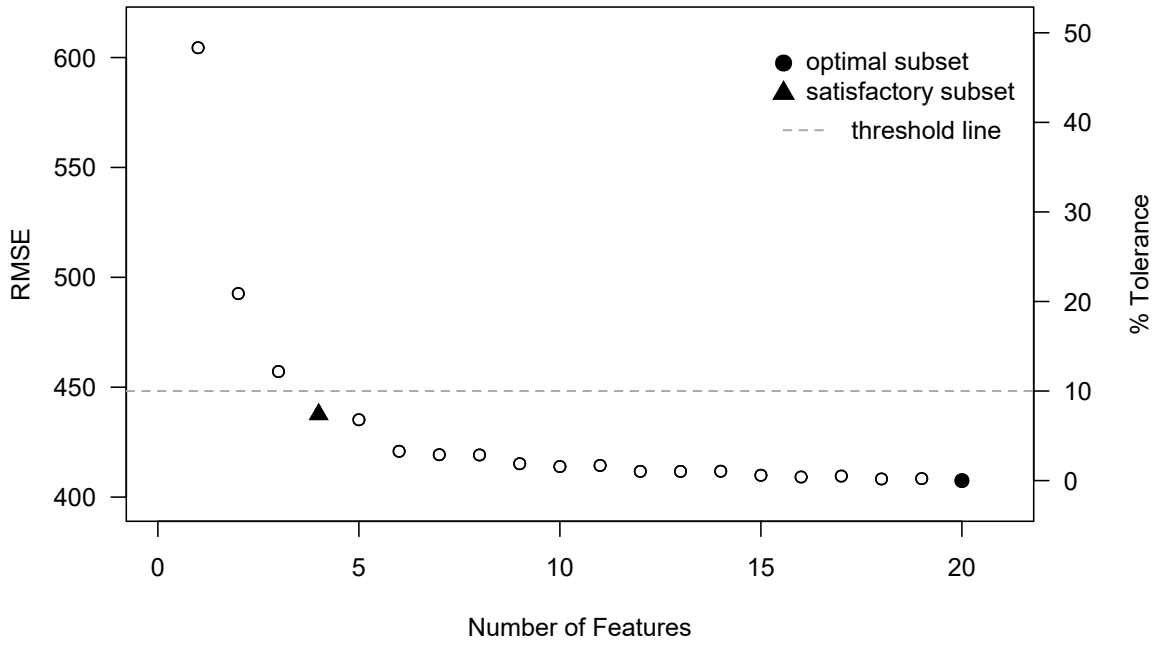



**Figure 2.6:** Feature Selection Protocol. Average RMSE by number of VIs (features). Optimal (filled circle) and satisfactory (filled triangle) subsets have 20 and four VIs, respectively. Optimal subset corresponds to the lowest average RMSE found through the feature selection process. The satisfactory subset corresponds to the minimal number of features (i.e., VIs) below a ten percent tolerance threshold over the minimal RMSE

### 2.3.3 Feature Filtering and Selection:

The filtering selection protocol reduced the number of features from 97 to 19 VIs. Within the selection workflow, the optimal (filled circle) and satisfactory (filled triangle) features subset were identified in the feature selection protocol (Figure 2.6). By this protocol, the number of VIs necessary to fulfill this constraint (i.e., 10% threshold above minimum RMSE) was equal to four and are listed in Table 2.3 as well as indicated in boldface in Table 2.2. Complementary, the optimal model, or the model with lowest RMSE value, includes all 20 VIs (i.e., filtered VIs plus the optimized NRI). However, the increments in performance are marginal from six features onward (Figure 2.6).

**Variable Importance**. The four selected VIs (satisfactory subset) used as features for different regression techniques are: the *optimized NRI*, *Chlorophyll Index*, *Simple Ratio 7* and *Double Difference near the red-edge*. Variable importance, extracted as per the description in Methods, respective formulas and bibliographical references are listed in Table 2.3.

**Table 2.3:** Selected Vegetation Indices (satisfactory subset), Variable Importance, Formulas and References.

| Vegetation Index | Importance | Formula | Author(Year) |
|---|---|---|---|
| Optimized NRI | 29.7 | $(\lambda_{755} - \lambda_{745}) / (\lambda_{755} + \lambda_{745})$ | Mutanga and Skidmore (2004a) |
| Chlorophyll Index | 27.5 | $(\lambda_{675} * \lambda_{690}) / \lambda_{683}^2$ | Zarco-Tejada et al. (2003) |
| Double Difference n | 17.2 | $2 * (\lambda_{710} - \lambda_{660} - \lambda_{760})$ | Lichtenthaler et al. (1996) |
| Simple Ratio 7 | 14.6 | $\lambda_{440}/\lambda_{690}$ | le Maire et al. (2008) |

### 2.3.4 Performance Analysis:

From the repeated $k$-fold cross-validation, the error metrics for each feature(s)-algorithm pair are summarized in Table 2.4. It presents, in descending order, the feature(s)-algorithm pair performances. Results range from (max) 770.6 to (min) 325.3 kg DM/ha (RMSE) from NDVI and SVM, respectively. In terms of precision, the *standard-deviation* (SD) ranges from 59.5 to 15.6 kg DM/ha (RMSE).

**Table 2.4:** Posterior analysis results for each feature(s)-algorithm pair.

| | RMSE | | | | $R^2$ | | | | MAE | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Models | Average | SD | Max | Min | Average | SD | Max | Min | Average | SD | Max | Min |
| NDVI | 636.9 | 59.5 | 770.6 | 532.1 | 0.49 | 0.05 | 0.61 | 0.39 | 495.0 | 38.1 | 567.5 | 430.0 |
| OptNRI | 520.3 | 51.0 | 692.2 | 406.2 | 0.65 | 0.08 | 0.77 | 0.45 | 412.0 | 36.2 | 514.1 | 311.1 |
| RPM | 521.8 | 58.8 | 645.5 | 374.8 | 0.68 | 0.07 | 0.82 | 0.54 | 406.6 | 45.8 | 511.8 | 296.8 |
| LM | 475.8 | 46.0 | 571.9 | 354.9 | 0.71 | 0.06 | 0.84 | 0.55 | 371.5 | 36.4 | 442.3 | 295.7 |
| MARS | 453.8 | 44.8 | 528.5 | 338.5 | 0.74 | 0.06 | 0.86 | 0.56 | 358.3 | 36.6 | 419.2 | 279.6 |
| SVM | 425.7 | 45.1 | 513.4 | 325.3 | 0.77 | 0.06 | 0.88 | 0.61 | 332.6 | 34.5 | 401.1 | 259.2 |
| RF | 407.6 | 41.7 | 493.5 | 327.6 | 0.79 | 0.05 | 0.88 | 0.63 | 312.3 | 33.1 | 386.7 | 248.9 |
| STACK | 405.8 | 15.6 | 434.2 | 367.5 | 0.79 | 0.02 | 0.82 | 0.74 | 310.7 | 12.5 | 337.8 | 278.3 |

The performance of RF and STACK models are equivalent in terms of average accuracy. However, the STACK model is more precise with a standard-deviation of 15.6 kg DM/ha in contrast to 41.7 kg DM/ha of the RF model. Such difference in error-distribution is statistically significant ($p$-value$<0.01$) as per the results of a two-sample KS-test results, displaying narrower error-distribution for the STACK model (Figure 2.7 - b).

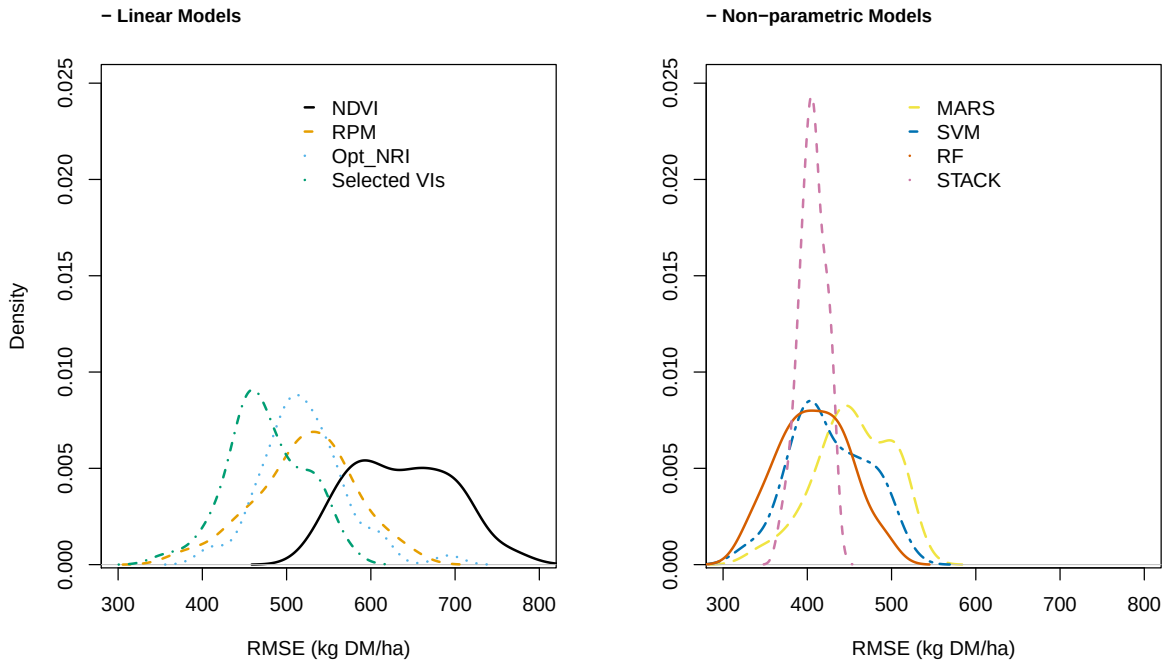



**Figure 2.7:** Density plot of the error-distribution RMSE for each model - (a): different features fitted to a linear model. (b): selected VIs (features) fitted using non-parametric methods.

Figure 2.7 presents the density plot of error-distribution of each feature(s)-algorithm pair, from which a visual assessment of accuracy and precision can be extracted. This figure is divided in to a density plot where only an ordinary linear regression (parametric) model was employed (Figure 2.7- a) and models where different non-parametric models and the selected VIs were employed (Figure 2.7 - b).

The visual analysis provided a clear indication of the higher precision of the STACK model in comparison with other methods. In the interest of avoiding similar figures, only the RMSE density plot will be presented as the same information for MAE and $R^2$ are also presented in Figure 2.8. The ranking and performance (average error metrics and error-distribution) of each feature(s)-algorithm pair is presented in Figure 2.8.

The results of the benchmark methodology developed by Kuhn and Johnson (2013) are presented in Table 2.5. With regard to accuracy differences, the best performing feature(s)-algorithm pair is the STACK of models with an average RMSE = 405.8 and MAE = 310.7 kg DM/ha. The lowest performing feature-algorithm pair is NDVI based on an ordinary linear model. Likewise, the saturation of NDVI becomes noticeable at around 2.500 kg DM/ha, whereas the optimized NRI does not present the same pattern (Figure 2.9). As a consequence, the maximum RMSE difference between feature(s)-algorithm pairs is between the STACK models and NDVI (231.1 kg DM/ha). Additionally, the optimized

NRI has, on average, a $\Delta$ (difference) in RMSE of 111.1 kg DM/ha smaller than the NDVI (Table 2.5).

Concerning error-distribution (precision), most noticeably the results of the $t$-tests (lower diagonal - Table 2.5), indicate that the average inaccuracy (RMSE and MAE) of (i) the STACK and RF models and (ii) RPM and optimized NRI are not statistically different (Table 2.5). Given such, it follows to check if each of these two pairs, (i) RPM/optimized NRI and (ii) RF/STACK, share equivalent error-distributions.

The $p$-value (0.717) for the two-sample KS-test indicates that the error-distribution between the RPM and the optimized NRI are not statistically different. These results indicates that both accuracy (Table 2.5 - $t$-test diagonal RMSE) and precision (KS-test) of the RPM and optimized NRI are equivalent.

In contrast, for the STACK and the RF algorithm, the KS-test $p$-value was 0.01. That indicates that both (error) distributions are unlikely ($p$-value $< 0.05$) to share the same distribution. In practical terms, these *post-hoc* tests indicate that the use of the STACK model provides more precise estimates, making it a more stable model, given that the standard-deviation of the error metrics (Table 2.4) for the STACK model is smaller than any other feature(s)-algorithm pair (Figure 2.8).

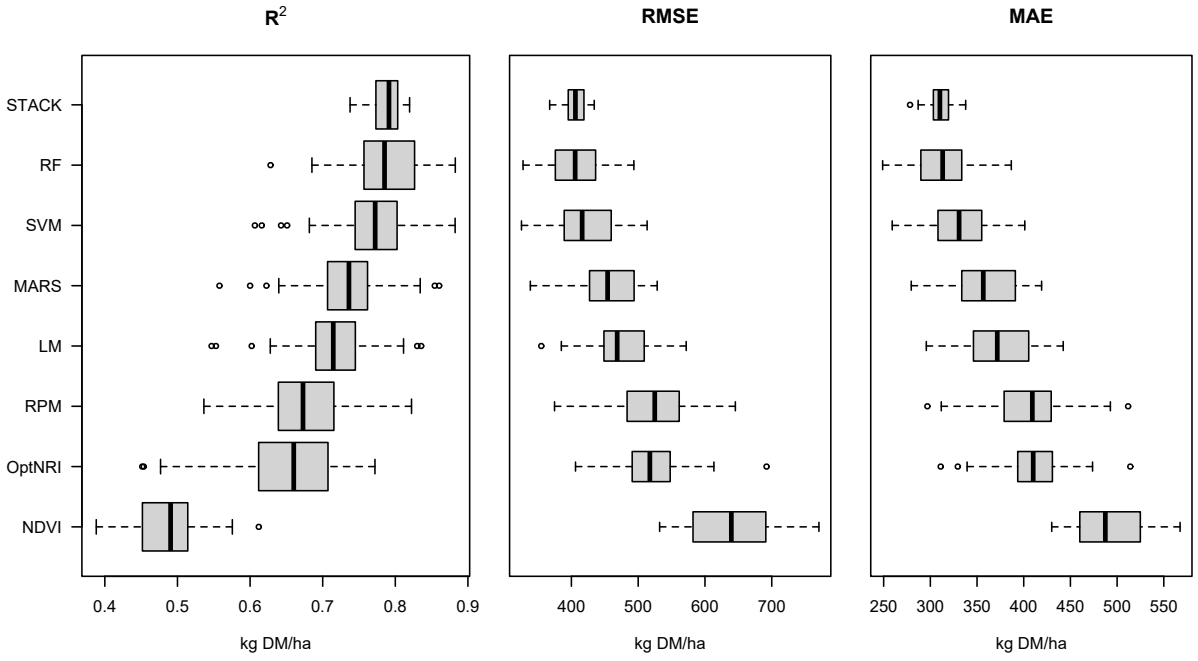



**Figure 2.8:** Error distribution presented through boxplots. The four top boxplots refer to non-parametric models. Conversely, the four boxplots on the bottom relate to ordinary linear models. The STACK model presents a narrower error-distribution.

**Table 2.5:** Posterior Analysis - RMSE, MAE and $R^2$ for feature(s)-algorithm pairs.

| Models | MARS | LM | SVM | RF | OptNRI | RPM | NDVI | STACK |
|---|---|---|---|---|---|---|---|---|
| ***MAE*** | | | | | | | | |
| MARS | | -25.10 | 11.36 | 31.52 | -69.12 | -59.11 | -147.45 | 36.83 |
| LM | 0.00 | | 36.46 | 56.61 | -44.02 | -34.01 | -122.35 | 61.92 |
| SVM | 0.01 | 0.00 | | 20.15 | -80.48 | -70.47 | -158.81 | 25.46 |
| RF | 0.00 | 0.00 | 0.00 | | -100.63 | -90.63 | -178.96 | 5.31 |
| OptNRI | 0.00 | 0.00 | 0.00 | 0.00 | | 10.01 | -78.33 | 105.94 |
| RPM | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | -88.34 | 95.94 |
| NDVI | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 184.27 |
| STACK | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | |
| ***RMSE*** | | | | | | | | |
| MARS | | -38.31 | 13.31 | 29.26 | -78.45 | -74.40 | -189.51 | 41.60 |
| LM | 0.00 | | 51.62 | 67.57 | -40.15 | -36.09 | -151.21 | 79.91 |
| SVM | 0.00 | 0.00 | | 15.95 | -91.76 | -87.71 | -202.82 | 28.29 |
| RF | 0.00 | 0.00 | 0.00 | | -107.72 | -103.66 | -218.78 | 12.34 |
| OptNRI | 0.00 | 0.00 | 0.00 | 0.00 | | 4.05 | -111.06 | 120.05 |
| RPM | 0.00 | 0.02 | 0.00 | 0.00 | 1.00 | | -115.11 | 116.00 |
| NDVI | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 231.11 |
| STACK | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | |
| $\boldsymbol{R^2}$ | | | | | | | | |
| MARS | | 0.04 | -0.02 | -0.03 | 0.10 | 0.06 | 0.25 | -0.05 |
| LM | 0.00 | | -0.06 | -0.08 | 0.05 | 0.02 | 0.21 | -0.09 |
| SVM | 0.00 | 0.00 | | -0.02 | 0.11 | 0.08 | 0.27 | -0.03 |
| RF | 0.00 | 0.00 | 0.01 | | 0.13 | 0.09 | 0.28 | -0.02 |
| OptNRI | 0.00 | 0.00 | 0.00 | 0.00 | | -0.04 | 0.15 | -0.15 |
| RPM | 0.00 | 1.00 | 0.00 | 0.00 | 0.24 | | 0.19 | -0.11 |
| NDVI | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | -0.30 |
| STACK | 0.00 | 0.00 | 0.01 | 1.00 | 0.00 | 0.00 | 0.00 | |

*Footnote* Upper diagonal: average difference between each feature(s)-algorithm pair (MAE, RMSE are expressed in kg DM/ha ). Lower diagonal: $t$-test $p$-value for H0: difference = 0.05. $p$-value adjustment: bonferroni

Validation results, in which trained-tested models are applied to an unseen (validation) dataset, are presented in Figure 2.9. The first row of plots presents predicted vs. observed results for different features in ordinary linear regressions (parametric-model). The second row presents the performance of selected VIs in different non-parametric regression algorithms.

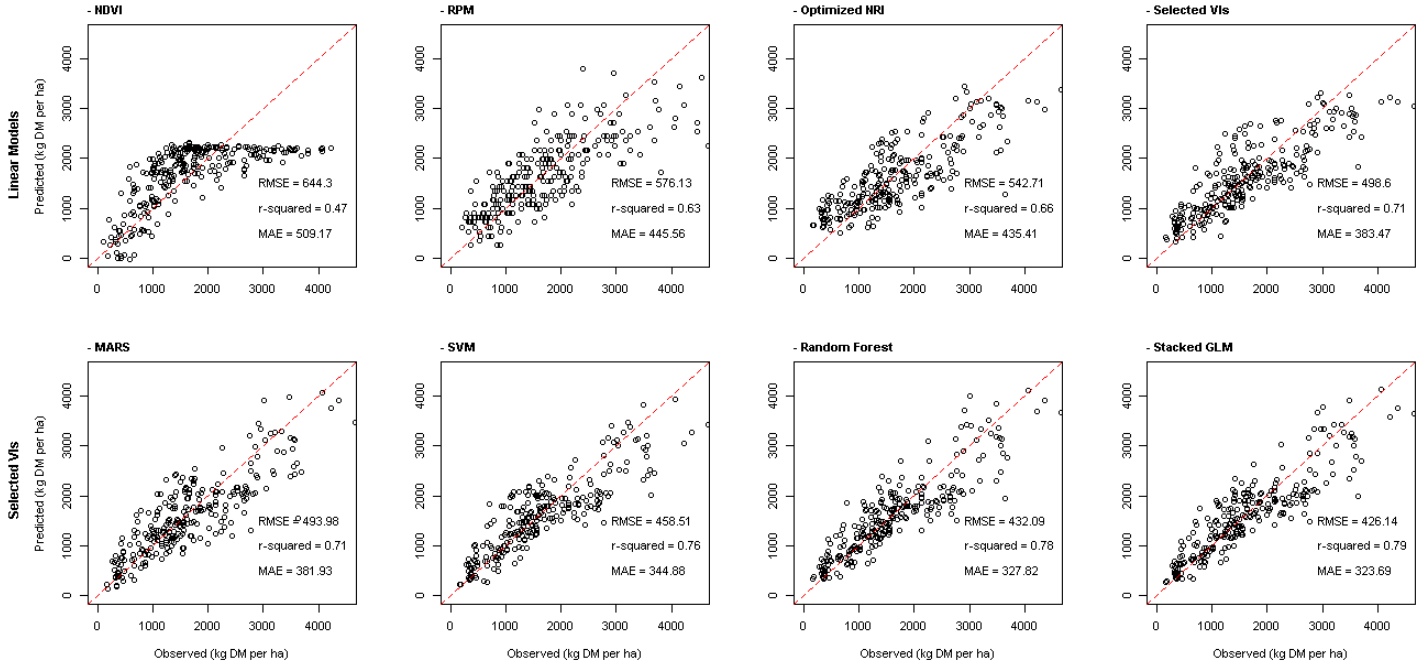



**Figure 2.9:** Scatterplots of predicted vs observed data of all feature(s)-algorithm pairs on the validation dataset ($n = 270$). The first row presents different features using an ordinary linear regression (parametric-regression). The second row uses the four selected VIs (*i.e., satisfactory subset*) in different (non-parametric) regression algorithms. Noticeably, the NDVI (top-left) scatterplot illustrates its saturation issue at around 2.500 kg DM/ha. Scatterplots have been arranged according to accuracy improvements and follow a left-to-right and top-to-bottom order.

## 2.4 Discussion

This study successfully selected and validated a subset of (four) VIs and regression technique (STACK) that largely outperforms the conventional method (RPM) for biomass estimation by an average $\Delta$ RMSE of 116kg DM/ha (Table 2.5).

This analysis has also shown that, in comparison to the most commonly adopted method on precision agriculture (i.e., standard NDVI), the best method substantially decreases the average inaccuracy and error-distribution (RMSE $\pm$SD) of biomass estimation from 637 $\pm$60 to 406 $\pm$16 kg DM/ha (Table 2.4), while eliminating its characteristic saturation issue (Figure 2.9 - NDVI). In retrospect, given the worse performance of NDVI in comparison with the RPM ($\Delta$ RMSE = 115 kg DM/ha), it is clear why end users have not adopted NDVI over RPM measurements.

The approach of an exhaustive-search of an optimized NRI identified a band combination that, both in terms of accuracy and precision, is equivalent to the RPM (Table 2.5). The performance of this particular VI (optimized NRI $\lambda_1 = 745$, $\lambda_2 = 755\,\text{nm}$) is in-line with the results reported by Mutanga and Skidmore (2004a). Equivalent performance and similar spectral band combinations were also presented in Cho et al. (2007), who reported that such combination overcomes the saturation issue presented by the standard NDVI (Figure 2.9 - optimized NRI).

An additional set of 97 other VIs was tested and compared based on a linear regression performance (Table 2.2): all were suboptimal in performance when compared to the optimized NRI. However, it is important to remark that the performances of these 97 VIs as predictors may be handicapped by a poor fit (Table 2.2), as the best fit function between a VI and biomass is not necessarily linear. Yet, the analysis demonstrated that the VIs extracted from the literature present a high degree of multicollinearity. The process of filtering eliminated 80.4% (78 out of 97) of features (VIs): a strong indication of redundancy and limitation of VIs. Also, the selection protocol indicated that models with more than four VIs only yield marginal accuracy improvements (Figure 2.6).

This analysis also quantified the improvement in error-metrics when employing non-parametric models in comparison to a parametric model (i.e., ordinary linear regression - LM). On average, the use of ordinary linear regression, with the same selected VIs, presented an additional 38.3, 51.6, 67.6 and 79.9 kg DM/ha error (RMSE) compared to MARS, SVM, RF and STACK models, respectively. It is important to highlight that such improvement in performance is due exclusively to model selection and tuning, improving results without requiring additional features.

While there are accuracy improvements based solely on the choice of employing either parametric or non-parametric models (e.g., STACK - LM = $\Delta$ 79.9 kg DM/ha), such improvements are not so substantial within the group of non-parametric models. The

largest difference between non-parametric models (RF - MARS) was equal to 29.3 kg DM/ha. Overall, the largest improvements in performance are due to a better selection of features (e.g., LM - NDVI = $\Delta$ -151.2) rather than a better performing regression algorithm. This highlights the importance of feature selection and feature engineering (e.g., optimized NRI process).

It was not surprising that both the RF and STACK techniques share the same average accuracy. Random forests employs the technique of bootstrapping the dataset and aggregating results (known as bagging). The same core concept is used in stacked models. However, STACK models present a better precision level, given that their output is an ensemble of other models (not features).

In essence, this analysis has shown that the use of a small feature set of VIs coupled with non-parametric methods enhances the accuracy and precision of pasture biomass estimation. Such VIs are located in the region of the red-edge and NIR-shoulder (Figure 2.10), regions which have been identified as important for chlorophyll and/or leaf area index (LAI) assessment (Darvishzadeh et al., 2008; Delegido et al., 2011) as well as for pasture biomass estimation (Clevers et al., 2007).

It is reasonable to hypothesize that the selection of VIs around the red-edge could be linked with an underlying physiological process as canopies with higher levels of biomass will present higher LAI and chlorophyll mass (drivers of shifts on the red-edge). Previous studies have reported similar findings: Horler et al. (1983) credits the red-edge shift, centered at around 740 nm, to leaf stacking, thus, showing a causal link with biomass. Tucker (1977b) and Collins (1978), indicated the same spectral region (740 nm) and behavior when estimating attributes linked to biomass. Guyot and Baret (1988) examined the effects of higher chlorophyll content and LAI on the spectral behavior around the red-edge. Finally, Mokvist et al. (2014) discussed a wider range of the limits of the far-red spectral absorption, demonstrating a plasticity for light absorption at higher biomass levels, indicating a characteristic which might overcome saturation.

Given the similar results and the mechanistic link explored by these previous studies, the approach and findings become more robust and, most likely, generalizable to other locations or contexts.

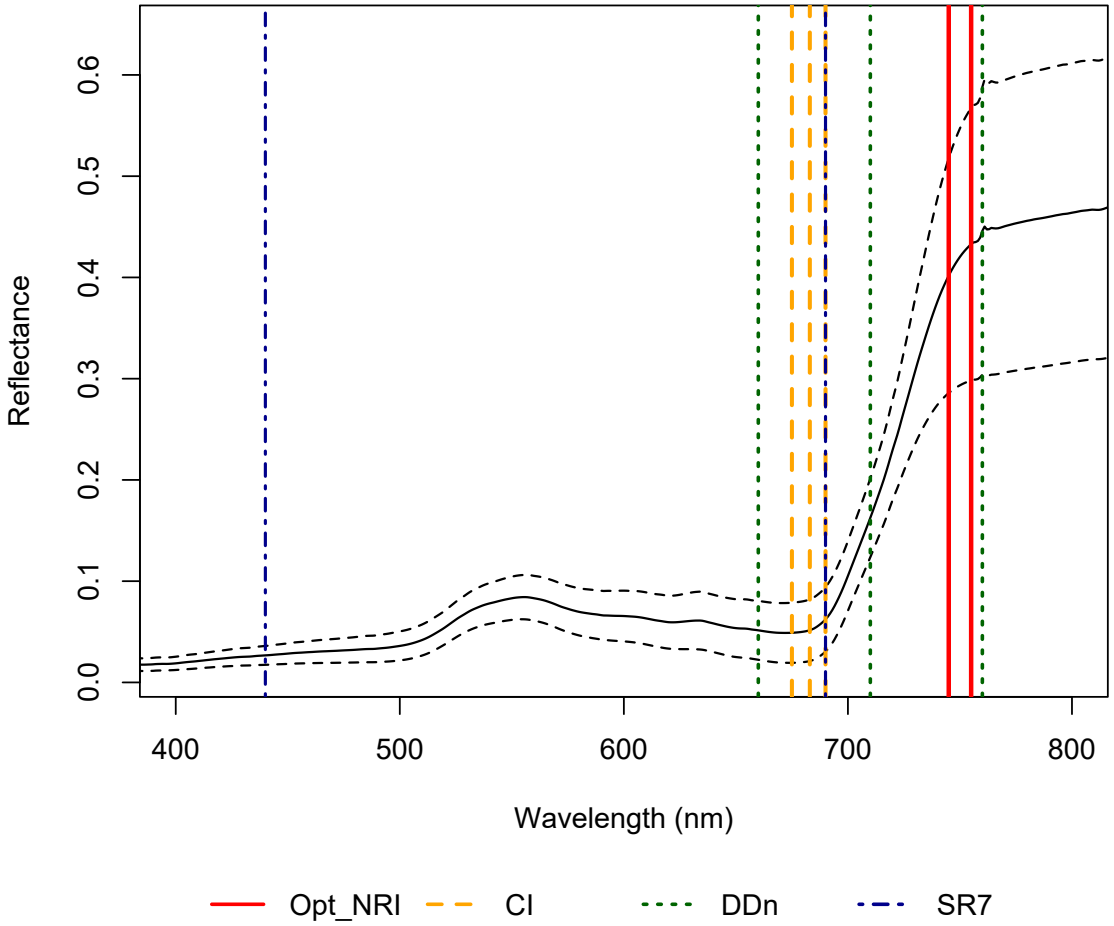



**Figure 2.10:** Selected VIs through the feature selection workflow. In the background, the mean (solid line), upper and lower boundaries (SD, dashed line) of all spectral observations. Different line styles identify each selected VI. All VIs have at least one band within the regions of chlorophyll absorption, the red edge and the NIR shoulder

The plateau of 430 kg DM/ha (RMSE) is equivalent to approximately 6.5 g per sample area (0.15 $m^2$). Thus, a fraction of this error may well be due to the sampling/harvesting technique or even a drying process that, despite best efforts, has an inherent random noise. It is reasonable, therefore, to assume that this methodology was able to explain most of the variance which could be mapped by VIs. Further performance improvements in error metrics are mostly likely due to overfit or noise-fitting. As illustrated in Figure 2.6, the addition of extra VIs will only yield marginal improvements. Under normal field-conditions, it seems unlikely that measurement error could be decreased further.

The ranges presented in Table 2.1 are in-line with growth rates, canopy architecture and height displayed in different seasons as per the intrinsic ecophysiological characteristics of

cool-season grasses (Christie et al., 2018). Given the significant duration and number of data collection campaigns, coupled with the consequent different plant development stages, this dataset should provide an adequate surrogate for on-farm pasture conditions.

As spectral data is highly correlated, it is unlikely that the relationships found would not perform equally well whenever employed in perennial ryegrass pastures displaying the same range of biophysical and biochemical status to the sample site in this study, granted that no major factor which could influence spectral response is introduced (e.g changing light/atmospheric conditions or soil-background at low biomass levels). This indicates a satisfactory approach to estimate a wide range of pasture biomass values throughout the year with a small number of VIs, which could be implemented on a UAV mounted sensor.

## 2.5 Conclusions

In this study, we examined the accuracy and precision of vegetation indices and machine learning methods for biomass assessment, while featuring the rising plate meter (RPM) as benchmark for pasture management. This study was able to present, test and evaluate an optimal method to select and employ spectral vegetation indices (VIs) for biomass assessment of perennial ryegrass under different fertilization, regrowth periods and seasonal effects. Findings are in-line with previous results reported in the literature, while our best method largely outperforms standard practices (i.e., RPM) as well as known VIs.

Our study has shown that the optimized NRI is equivalent to the RPM while providing a firm analytical ground for substitution between these techniques. Also, it is clear that reflectance can outperform canopy height as a feature for biomass estimation provided that the optimal spectral bands and methods are employed. Furthermore, multispectral systems can be employed in a multitude of data acquisition platforms, providing estimates at near or real-time.
This analysis could lead to the development of a multispectral UAV-mounted sensor with known added-value in relation to current on-farm practices and precision agriculture methods. Overall, the comparison of achievable accuracies by employing either canopy height or reflectance (through VIs) can indicate the most appropriated path when developing sensors and establishing optimal sensing techniques (either spectral or canopy based) for assessing pasture biomass.

Given the agreement of these findings with previous research, the best method (STACK and selected VIs, as feature(s)-algorithm pair) and validation results (i.e., assessed against an unseen validation set) should be transferable from this study site to other locations. Furthermore, from an end-user perspective, the accuracy and precision achieved is sufficient to efficient pasture-management.

Given the high accuracy achieved, further improvements to this methodology will necessarily have to take into account the accuracy of reference measurements and data-collection protocols. However, given the unsatisfactory results of current VI methods, a new UAV sensor (with a small number of bands) should benefit from the results and methods presented here. Finally, this study also indicated that the most well-known VI (NDVI) is not optimal for biomass estimation. It also indicated that better performing VIs which could be measured from a UAV-mounted multispectral sensor, which would allow mapping and monitoring of pasture biomass with high accuracy at very high spatial resolution and with complete coverage of management areas.

# Chapter 3

# Retrieval of Crude Protein in Perennial Ryegrass Using Spectral Data at the Canopy Level

This chapter is based on:

G. Togeiro de Alckmin, A. Lucieer, G. Roerink, R. Rawnsley, I. Hoving, and L. Kooistra (2020b). "Retrieval of Crude Protein in Perennial Ryegrass Using Spectral Data at the Canopy Level". *Remote Sensing* 12.18, 2958. DOI: 10.3390/rs12182958. URL: https://www.mdpi.com/2072-4292/12/18/2958

*Supplementary materials to this chapter can be found in the online publication.*

# Abstract

Crude protein estimation is an important parameter for perennial ryegrass (*Lolium perenne*) management. This study aims to establish an effective and affordable approach for a non-destructive, near-real-time, crude protein retrieval, based solely on top of canopy reflectance. The study contrasts different spectral ranges, while selecting a minimal number of bands and analyzing achievable accuracies for crude protein expressed as a dry matter fraction or in a weight per area basis. In addition, model prediction performance in known and new locations are compared. Data collection comprises 266 full-range (350 - 2500 nm) proximal spectral measurements and corresponding ground truth observations in Australia and the Netherlands from May to November 2018. An exhaustive-search (based on a genetic algorithm) successfully selected band subsets within different regions and across the full spectral range, minimizing both number of bands and an error metric. For field conditions, our results indicate that the best approach for crude protein estimation relies on the use of the visible to near-infrared range (400 - 1100 nm). Within this range, eleven sparse broad bands (of 10 nm bandwidth) provide better or equivalent performance than previous studies which used a higher number of bands and of narrower bandwidth. Additionally, when using top of canopy reflectance, our results demonstrate that the highest accuracy is achievable when estimating crude protein in its weight per area basis (RMSEP 80 kg/ha). These models can be employed to new unseen locations results with a minor decrease in accuracy (RMSEP 85.5 kg/ha). Crude protein as a dry matter fraction presents a bottom-line accuracy (RMSEP) ranging from 2.5 - 3.0 %DM in optimal models (requiring ten bands). However, these models display a low explanatory ability of the observed variability ($R^2 > 0.5$), rendering it only suitable for qualitative grading.

## 3.1 Introduction

Timely assessment of feed quality parameters is necessary for optimal management of pasture-based dairy (livestock) systems. Such monitoring can instruct well-informed decisions in grazing (feeding) management, according to whichever nutritional goals are established (Fariña et al., 2011). Currently, these parameters are usually estimated through laboratory analysis, involving time-consuming sampling procedures, complex logistical operations and are subject to the availability of a service provider (Jones et al., 1987; Roberts et al., 2004). In practice, most pasture managers often rely on proxies such as seasonal patterns, plant maturity, morphology or ecophysiological relations (Prewer et al., 2004), which may not provide the accuracy nor precision required for optimal pasture management (Machado et al., 2005). Within perennial ryegrass' (*Lolium perenne*) macronutrients, *crude protein* (CP) usually displays the largest variability (Roberts et al., 2004), directly responding to soil-plant-animal interactions and seasonal patterns, whilst being a key component to optimal ruminal activity and, consequently, intake and digestibility (Valk et al., 2000).

Feed quality assessment is ultimately based on wet-chemistry methods which are usually expensive, complex and time-consuming. Inversely, *near-infrared spectroscopy* (NIRS) is an established spectral-based laboratory technique able to reduce costs and increase throughput, while providing estimates equivalent to such reference methods (Jones et al., 1987; Norris et al., 1976). Ideally, employing non-destructive spectral analysis in field conditions would be advantageous to farmers as accurate real-time information allows precise management of the herd's diet and understanding of short and long-term effects of different strategies for grazing (e.g., grazing interval and pressure) and pasture management (e.g., timing and application rates of inputs).

To such end, *remote sensing* (RS) has been employed in several studies (Curran, 1989; Kawamura et al., 2008; Sanches et al., 2013; Wijesingha et al., 2020), where different spatial, spectral and data acquisition scales (e.g., proximal, low-level flight, airborne and satellite) were examined. None, however, has found the same level of accuracy for crude protein retrieval as benchtop NIRS, demonstrating the challenging nature of the straightforward translation of laboratory methods to field conditions. In contrast to processed laboratory samples (i.e., dried and ground), spectral field measurements are impaired by non-homogeneous biophysical attributes of targets (e.g., moisture, *LAI*, *LAD*) (Asner, 1998), in which the underlying fundamentals of spectroscopy (Beer-Lamberts' law) (Beer, 1852; Curran, 1989) are not entirely observed.

In outdoor environments, measurements and analyses are increasingly complex given the dynamic and reactive nature of plant-light interactions and nitrogen mobility through the plant (Baranoski and Eng, 2007; Shorten et al., 2019). More importantly, field spectroscopy is an ill-posed problem as per Hadamards's definition (Hadamard, 1902): an observed

spectral response can be resolved by multiple combinations of parameters (e.g., pigment concentration, optical thickness), resulting in ambiguous estimations of the target's true constituents (Asner, 1998; Baret and Samuel, 2008). Consequently, in the absence of prior biophysical information, plant-tissue biochemical characterization based solely on *top of canopy* (TOC) reflectance is inherently poorly determined.

To cope with such limitations, two approaches have been extensively utilized (Baret et al., 2006; Biewer et al., 2009; Kawamura et al., 2010; Starks and Brown, 2010; Thulin et al., 2014): (i) the use of spectral shapes and transformations (e.g., derivatives and continuum-removal) and (ii) the estimation of the canopy's biochemical attributes in conjunction with biophysical properties (e.g., biomass per unit area). Such approach (ii) has led to the retrieval of biochemical estimates as mass per unit area (i.e., kg/ha) rather than as a *dry matter* (DM) fraction (i.e., %DM) (Baret and Fourty, 1997; Baret et al., 2006).

While spectral transformations (i) are used as a tool to disentangle overlapping absorption features and ameliorate light scattering, in practice, such potential has not been confirmed in uncontrolled environments when employed as a preprocessing tool to improve estimations (Starks et al., 2006; Thulin et al., 2014). Additionally, in a farm-scenario, its adoption would be limited by the necessary number of contiguous bands, narrow-bandwidth, radiometric sensitivity and high signal-to-noise ratio necessary for its accurate computation. In contrast to transformations, reflectance measurements can be executed in sparse bands, with broader bandwidths and coarser radiometric resolution. Hence, a plausible alternative is the (ii) estimation of biochemical attributes in terms of mass per area through reflectance measurements (Kattenborn et al., 2019).

Despite not achieving the same levels of accuracy as NIRS, from a pasture management perspective, the absence of extremely accurate estimates (i.e., within less than 1% of the reference method) is offset by continuous non-destructive estimations of acceptable precision, provided that these are made available at a low-cost and without geographical or temporal restrictions. These requirements can be met through the development of a spectral model with a small number of bands (i.e., feature selection), validated under representative sampling conditions (i.e., different locations, dates or seasons). Such conditions are essential to instruct the design of an effective, low-cost sensor for CP estimation.

Notably, hyperspectral data presents a high degree of multicollinearity (i.e., redundancy of explanatory variables), which is usually handled through multivariate techniques such as *partial least squares regression* (PLSR). However, a common misconception is the assumption that the full spectral range should yield better results than the causal or informative part of the spectra (Kjeldahl and Bro, 2010), allowing for a feature selection routine without a corresponding loss on model prediction accuracy (Rasmussen and Bro, 2012). However, due to overlapping absorption features and multicollinearity, an upfront and deliberate choice of optimal spectral ranges or band subsets (i.e., explanatory

variables) is not trivial. Multicollinearity also renders model comparison and performance analysis difficult as different subsets of bands can be employed with equivalent model performance.

Approaching these issues, Kawamura et al. (2010) provides an insight into feature selection for *crude protein yield* ($CP_m$, kg CP/ha) and achievable accuracies employing both iterative stepwise elimination and *genetic algorithm* (GA) routines, reporting best results in the latter ($R^2 = 0.79$ and 0.85, respectively). The authors also applied the same routines to estimate %CP and, as expected, reported lower accuracies than $CP_m$ ($R^2 = 0.30$ and 0.27, respectively).

To a large extent the cost-requisite can be tackled by a restriction of spectral regions employed in models and, consequently, sensors. *Visible and near-infrared* (VIS-NIR, 400–1100 nm), mostly *Si*-based semiconductors, mass-produced for off-the-shelf consumer cameras, are less expensive than *shortwave infrared* (SWIR, 1100 - 2500 nm) instruments, mostly *InGaAs* or *PbS*-based. Discussing this issue, Starks et al. (2006) reported that CP (%CP and $CP_m$) could be best estimated using the VIS-NIR rather than the SWIR portion of the spectrum.

By joining these methods (Kawamura et al., 2010; Starks et al., 2006), this study aims to assess model performances coupled with routines of feature selection within different spectral regions (i.e., VIS-NIR or SWIR) or the full spectrum (FS) range. Such analysis can inform both the minimal necessary spectral measurements for crude protein retrieval as well as the design of a *multispectral* (MS) sensor (both point-measurement or imaging system), less complex and costly than hyperspectral instruments employed in laboratory analysis.

In summary, this research aims to predict perennial ryegrass $CP_m$ and %CP after routines of feature selection (i.e., genetic algorithm) employing either VIS-NIR, SWIR or the full-range spectrum as predictors. Spectral transformation (e.g. derivatives or continuum-removed features) are not to be employed, as these require either contiguous or specific bands and low radiometric error, which may prevent the transferability of findings to a rural scenario. Finally, the analysis of the data should indicate a necessary minimal number of spectral bands for the design of a multispectral sensor at low-cost with an acceptable accuracy, which can be integrated in an autonomous system (unmanned aerial or ground vehicle) for a higher level of automation in data collection.

## 3.2 Methods

The experimental units consisted of rainfed perennial ryegrass plots managed under different *nitrogen* (N) fertilization regimes and mowing intervals, generating a quantitative and qualitative gradient. The combination of these factors (i.e., nitrogen rate and regrowth interval) was randomized across each *pseudo*-replicate (Figure 3.1 - A and B). Nitrogen

(supplied as urea) was broadcast uniformly across each plot. For all plots, the residual mowing-height was approximately 50 mm (Figure 3.1 - II).

Four experimental sites were available: Elliot, Goutum, Vredepeel, and Zegveld. While the first is located in Australia (Tasmania), the remaining sites are in the Netherlands. The data collection period started in early May and finished in late November 2018 (Table 3.1).

**Table 3.1:** Metadata associated with the data collection per location.

| Location | Date 1 | Date 2 | Date 3 | Instrument | Soil Type | N Rates * | Regrowth ** |
|---|---|---|---|---|---|---|---|
| Elliot (AU) | 11-Nov | 17-Nov | 24-Nov | FieldSpec 4 | Clay | 50-75-100 | 3 |
| Goutum (NL) | 15-May | 28-Jun | 05-Oct | FieldSpec 3 | Clay | 180-360 | 4 |
| Vredepeel(NL) | 11-Jun | 03-Jul | 27-Sep | FieldSpec 3 | Sandy | 180-360 | 4 |
| Zegveld (NL) | 10-May | 19-Jun | 10-Oct | FieldSpec 3 | Peat | 180-360 | 4 |

* kg N/ha ** Weeks/Mowing Cycles

### 3.2.1 Experimental Setup

The Dutch experimental sites (Figure 3.1 - A) contained 24 plots per location. In each, the layout was a factorial combination of three N fertilization levels (0, 180, and 360 kg N/ha per year), four mowing intervals (in a cycle of four weeks, when then all plots were mown) and two *pseudo*-replicates of this combination, as presented in Hoving et al. (2018). Each site is located within a different soil type (either clay, sandy or peat). Data collection spanned from May to October 2018. From the second half of July till September, data collection was interrupted due to a prolonged heatwave and drought, which constrained plant growth.

The Australian site (Figure 3.1 - B) consisted of 30 plots, managed under five different N levels (0, 25, 50, 75, and 100 kg N/ha) applied on October 20$^{th}$, at a clay soil location. A set of 10 distinct plots was mowed on November 11$^{th}$ and 17$^{th}$, resulting in three different regrowth periods. Data collection took place in the 11$^{th}$, 17$^{th}$, and 24$^{th}$ of November, 2018.

To provide evidence that these models are not restricted by location or season, data was collected in four locations (one in Australia and three in the Netherlands) spanning from Spring to Autumn (i.e., the most important growth period), during early May to late November 2018. Nitrogen levels were set based on previous research work and in accordance to appropriate ranges of fertilizer levels commonly applied in each production system (Hoving et al., 2018; Rawnsley et al., 2014; Valk et al., 2000). Dates were chosen to portray the most important growth periods (i.e., Spring and Autumn) for cool-season grasses. Only in Australia all data was collected in Spring (November), as the instrument was available for a limited period of time. However, Australian biomass and %CP ranges

were comparable to the locations in the Netherlands and analyzed within the *Ground Truth Analysis*.

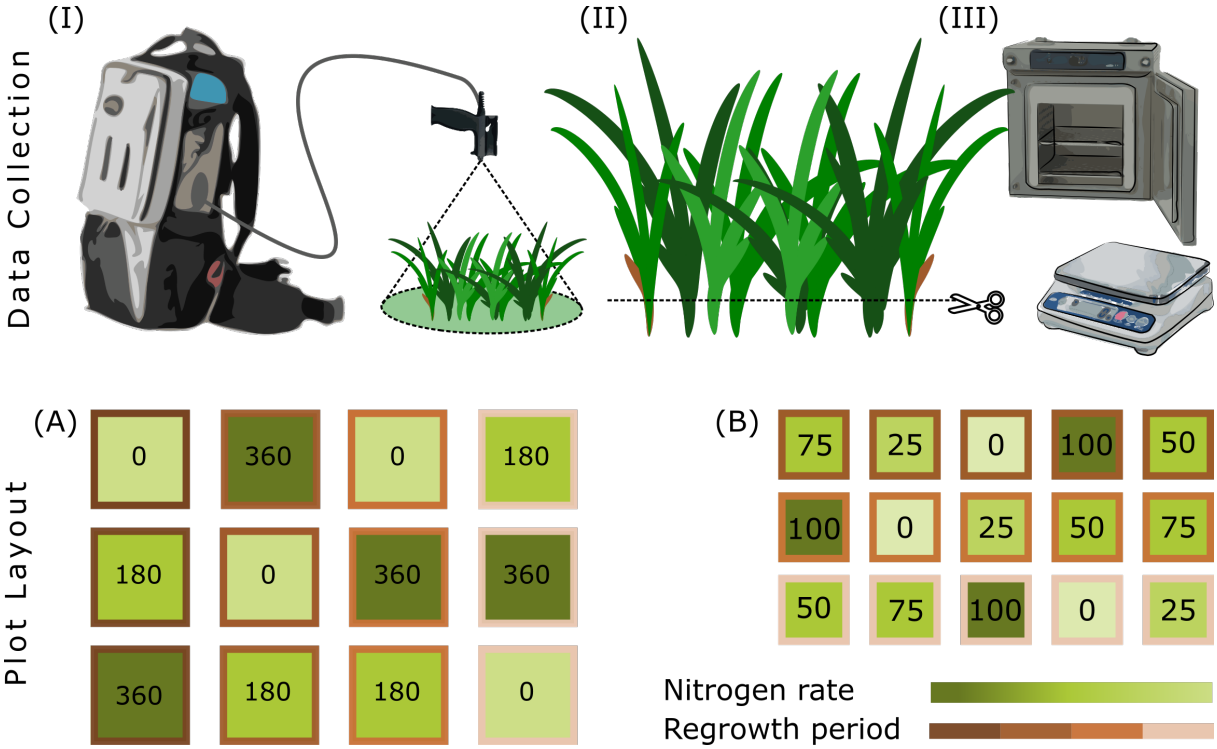



**Figure 3.1:** Data collection protocol (I - III) and plot layout (basic *pseudo*–replicate, A and B). Top: (I) Spectral Measurements, (II) Mechanical Defoliation and (III) Drying and Weighing. Bottom: basic Dutch (A) and Australian (B) plot layout. Borderline colours indicate regrowth period and hues of green (darker higher rates) indicate nitrogen levels (rates are also indicated within each plot)

### 3.2.2 Data Collection

An identical data collection protocol was performed across sites and dates (Figure 3.1 - I to III). Spectral measurements, ranging from 350 to 2500 nm, were taken in clear-sky periods at around solar noon ± 2 hours. Total time employed for spectral data collection ranged from 45 minutes to one hour. An ASD FieldSpec® 3 and FieldSpec® 4 (Malvern PanAnalytical - Colorado, USA) were employed in the Netherlands and Australia, respectively. Both instruments had no attached fore optics (i.e., bare fiber. Field of view: 25°) and were configured using the following setup (i.e., number of scans): 60 for white reference (Spectralon$^{TM}$. Labsphere - New Hampshire, USA), 60 for dark-current, 30 per measurement and five measurements of each target (Figure 3.1 - I). The average of these five measurements was utilized, for analysis purposes, as the target's reflectance value. Measurements were taken at approximately one meter high (Figure 3.1 - I), resulting in a circular footprint of approximately 0.44 m diameter.

After the measurements of each plot were taken, a reference scan of the white reference was recorded for quality assurance. This procedure aimed to create a *pseudo* dual field of view (Mac Arthur and Robinson, 2015), by which the target's reflectance could be further corrected in case of changing ambient light (e.g., cloud formation). The instrument was recalibrated whenever the white reference measurement deviated from a straight line centered at one or a maximum time-limit of seven minutes between recalibrations was reached, whichever occurred first.

The footprint was mechanically defoliated to a specific residual height (i.e., 50 mm) and stored in perforated plastic oven bags (Figure 3.1 - II). Such residual height corresponds to the best practice perennial ryegrass management (Fariña et al., 2011) as well as a height strata that has little to no light interaction (Lantinga et al., 1999). The harvested material was immediately refrigerated and transported from the experimental sites to a forced-air oven, where it was dried for 48 hours at 65 °C and weighed (Figure 3.1 - III).

The samples were then taken to a third-party feed quality laboratory (EuroFins Agro, NL and FeedTest, AU), where %CP was estimated using proprietary NIRS methods. Crude protein fraction estimates are expected to be within less than 1.0% RMSE (Marten et al., 1989) of wet-chemistry methods results.

Observations with sample-weight above 3500 kg DM/ha were excluded from the data set as these are not biomass levels to be achieved within the best pasture management practices (Fariña et al., 2011) (target pre- and post-grazing biomass of 2400 to 1500 kg DM/ha, respectively) and, thus, should not be considered within the representative conditions under which these models should operate.

### 3.2.3 Data Analysis

Data analysis was performed in RStudio/R (R Core Team and R Development Core Team (2019), versions 1.2.5 and 4.0.2, respectively). The main necessary packages for the analysis, besides the base and dependencies packages, are *rsample* (Kuhn and Johnson, 2019), *desirability* (Kuhn, 2016), *plsVarSel* (Mehmood et al., 2012), and *caret* (Kuhn, 2008). For reproducibility purposes, data analysis operations are introduced by the corresponding *package::function* format (italics typeface and accompanied by the double colon operator, i.e., the *scope resolution operator*).

*Ground Truth Analysis*

**Comparability between locations**: measurements underwent a series of *post hoc* tests (i.e., analysis of variance) to ensure that data collected between locations was comparable, testing whether average values($\mu$) and variances ($\sigma^2$) were significantly different ($H_0$: $\mu_i = \mu_j$ at $\alpha = 0.05$). Initially, the Bartlett Test of Homogeneity of Variances (*stats::batlett.test*) was performed, followed by Kruskal-Wallis (*stats::kruskal.test*) and Dunn's Multiple

Comparisons (*FSA::dunnTest*). The *Kruskal-Wallis* (KW) test is a non-parametric method for testing whether samples originate from the same distribution. Pearson-correlation ($\rho$, and its $p$-value *stats::cor.test*) between biochemical and biophysical attributes were also computed to identify spurious relationships between spectral based models and the main biophysical property (i.e., biomass). Such should indicate whether $CP_m$ models were estimating simply biomass or, indeed, the attribute of interest (i.e., $CP_m$).

**Missing data and outlier detection**: two samples had %CP missing values, as these did not reach the minimum sample weight required for testing. For these, %CP values were imputed (*missMDA::imputePCA*) using the methods described in Josse and Husson (2016), having spectral data and metadata (e.g. collection date, location) as predictors. Crude protein yield ($CP_m$) was calculated as the product of %CP per the sample's dry matter weight. Outlier detection was performed (*enpls::enpls.od*) as per the workflow described in Cao et al. (2017). The top ten observations that accounted for the highest error standard-deviation or error-mean were excluded from the analysis data set. Such process is equivalent to other preliminary analysis as the inspection of the inner relations of PLS (i.e., TU plot) where the scores of both the predictor ($T$) and response variables ($U$) are plotted in order to identify outliers or the analysis of Q-residuals and Hotelling Distance (Hotelling, 1931).

*Spectral Analysis*

**A. Preprocessing**: raw reflectance measurements were exported to ASCII format using ViewSpecPro™ and merged (*dplyr::right_join*) with the sample's metadata and feed quality information. The raw target-reflectances were summarized as the average of the five measurements and a data frame with single reflectance value per measurement was created. This data frame was transformed into a Speclib object (*hsdar*, Lehnert et al. (2018)), spectra was corrected at 1000 and 1800 nm (*prospectr::spliceCorrection* - FieldSpec sensor artifacts), followed by a smoothing process using a Savitzky-Golay filter (*hsdar::noiseFiltering*, window-size = 19 nm) and resampling (*hsdar::spectralResampling*) to 10 nm (FWHM) contiguous bands, using a gaussian distribution as response function. The Savitzky-Golay filter was set in such a way to remove high-frequency noise, prevent over-smoothing and windows-size is similar to the value (i.e., 15 nm) employed in Kawamura et al. (2010). Bands associated with water vapor absorption features and noisy areas were identified and excluded (i.e., 350 to 400, 1325 to 1475, 1775 to 2000 and 2400 to 2500 nm) from the final data frame (Figure 3.2 - Pre-processing). For visualization and reference purposes, an exploratory analysis was performed by grouping and averaging the spectra according to quartiles (either $CP_m$ or %CP).

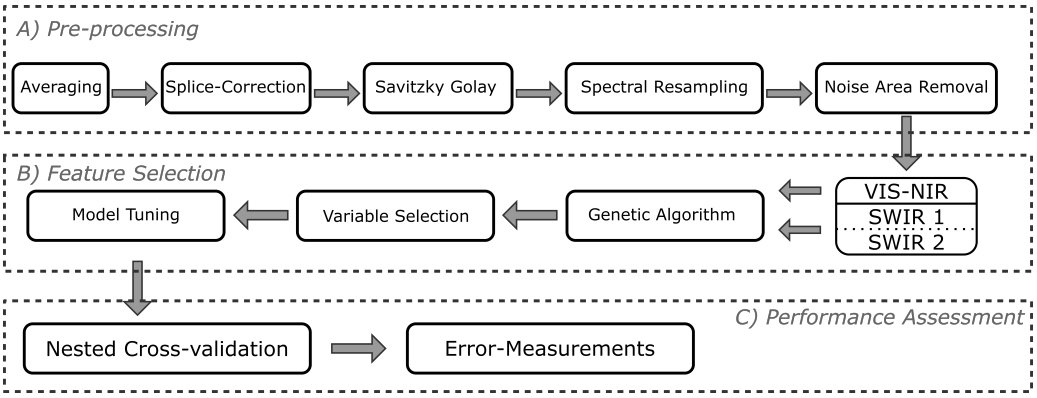



**Figure 3.2:** Data Analysis Workflow. The workflow is divided in (A) *Pre-processing*, (B) *Feature Seleection* and (C) *Performance Assessment*. Within the feature selection, three sets (VIS-NIR, SWIR, and FS) were analyzed in parallel for further performance assessments

**B. Feature (Bands) Selection**: the workflow for feature selection is described in Kuhn and Johnson (2013, ch. 19). A *genetic algorithm* (GA) routine for feature (i.e., spectral bands) selection (*caret::gafs*) coupled with PLSR (*pls::plsr*, Mevik and Wehrens (2015)) was carried out using the full spectrum, VIS-NIR (400–1100 nm bands) and SWIR (remaining bands from 1100–2500 nm nm). The genetic algorithm conducts a search procedure based on an initial number (i.e., population) of combinations of features (i.e., individuals). The search procedure aims to maximize an internal optimization function (Figure 3.3, *desirability*, Kuhn (2016)) that conjugates both an error metric (i.e., RMSE) and number of features. A response-surface for the internal maximization (objective) function is presented in Figure 3.3. For each spectral range, the function's parameters (i.e., number of bands and RMSE) were set in a way that its range spanned across possible solution-spaces for both initial and final generations. These ranges were found through an *unconstrained* model (*pls::plsr*), providing insight for achievable RMSE values. Otherwise, the GA could possibly be constrained in solutions-spaces where changes in performance would not necessarily equate to improvements in the objective function.

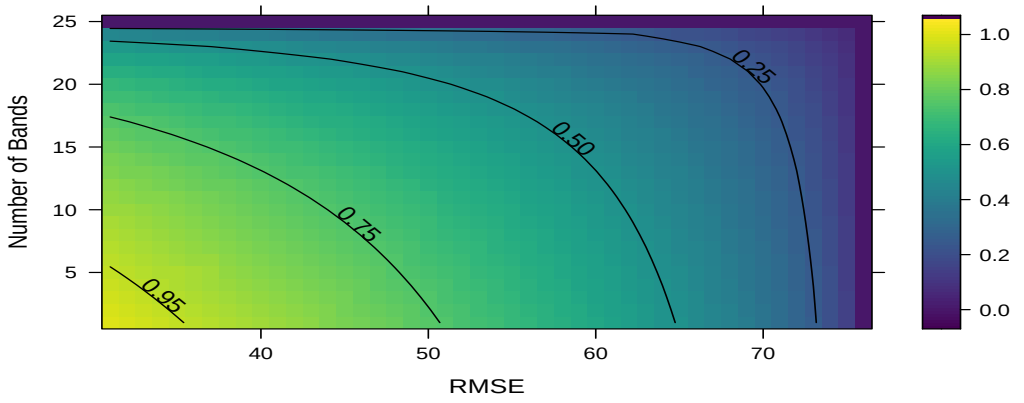



**Figure 3.3:** Internal Maximization Function Surface and Countour plot (`package::desirability`). The color gradient and contour lines indicates the response function of both model number of features and RMSE. While RMSE has a linear component, the number of bands (features) was set to a quadratic function.

Best subsets of the feature space were found through the crossover of populations, mutation of individuals and elitism, as described in in Scrucca (2013). A maximum number of generations (i.e., iterations) was set to 150 and the population size was set to 50. The range of 150 generations was set to provide an exhaustive search and, consequently, avoid either local minima or an early interruption of the search process. The population value of 50 was set to ensure a large population, in which the algorithm could best randomly select different subset of bands (Scrucca, 2013). To reduce the computational load, a maximum number of latent components was set to eight during feature selection. Such number of latent variables was chosen so at least 95-98% of the variance of the X matrix was utilized, ensuring comparison between models.

Each generation evaluates the overall fitness and selects the best individuals to be incorporated (as parents) to the next selection round (Figure 3.4). To assess possible overfitting, at each generation, an external assessment (i.e., RMSE) of the best individual was also computed. Such assessment is linked to a 10-fold cross-validation where, at each fold, 90% of data was used to train and 10% was held out as an external assessment. Consequently, the GA routine coupled with the internal maximization function was carried out 10 times for each spectral range and independent variable (i.e. either $CP_m$ or %CP).

A final selection procedure, using all 10-fold's best individuals as initial population, was carried out against the complete data set. The minimal RMSE individual is considered the optimal feature subset. In total, for each of the feature selection processes, the number of PLSR models (i.e., individuals) created was equal to 50 * 1500 as a result of the population

size (50), the total number of generations (150), and the cross-validation process (10-fold cross-validation).

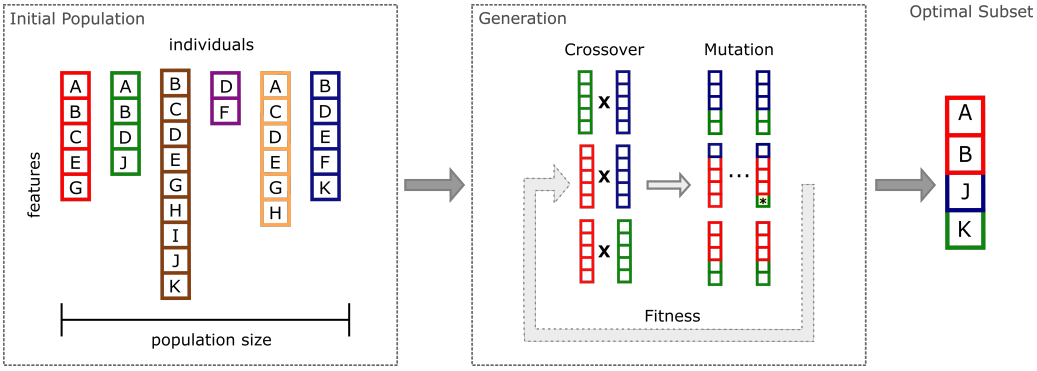



**Figure 3.4:** Genetic algorithm workflow. An initial set of candidate solutions is created from randomly selected features. These candidates are further selected based on a internal fitness measurements. As result, an optimal subset is identified. Tuning of the algorithm is done through adjustment on the population size, levels of crossover and mutation, number of generations and fitness measurements.

*Variable Importance*: was assessed through two different parameters: *variable importance in projection* (VIP - Equation 3.1), and *significance multivariate correlation with PLS* (sMC - Equation 3.2), which were calculated according to Mehmood et al. (2012) and the associated R-package *plsVarSel*.

$$v_j = \sqrt{p \sum_{a=1}^{A} (SSY_{comp,a}) \times W_{a,j}^2) / SSY_{cum}} \tag{3.1}$$

$$sMC_j = \frac{MS_{j,PLS_{regression}}}{MS_{j,PLS_{residuals}}} \tag{3.2}$$

where:

$p$ : number of variables;

$v_{\mathbf{j}}$ : variable importance in projection of the j$^{\text{th}}$ feature (band);

$a$ *to* $A$ : a$^{\text{th}}$ component to maximum (A) number of latent components;

$j$ : the j$^{\text{th}}$ feature;

$SSY$ : sum of squares of Y;

$W$ : loading weights of the predictor matrix (i.e., features);

$MS$ : mean squares (from the $F$-test notation).

VIP calculates the cumulative measure of the influence of individual features (variables) on the model. As a general rule, features with a VIP lower than 1 are considered not important.

The sMC (Equation 3.2) returns the $F$-test value of each predictor, while also calculating the significance $F$-value threshold (at $\alpha$ =0.05). Features with $F$-test value higher than the threshold level are considered not important.

**C. Performance Assessment and Validation**: nested cross-validation was chosen to assess model tuning and performance. In it, hyperparameter optimization (i.e., number of latent components) is performed within the inner cross-validation (i.e., *bootstrap*), while the outer cross-validation computes an unbiased estimate of the expected model's average error and distribution. The outer validation, however, is performed using two different strategies: (I) a $k$-fold cross-validation (*rsample::v_foldcv*) and (II) a group $k$-fold cross-validation (*rsample::group_vfold_cv*). While the first was assigned to randomly choose measurements from all experimental sites, the latter held out a single location at a time as the validation set. These two strategies should allow model comparison performance at locations with *prior information* and at *unseen locations* (Figure 3.5). The first strategy (Figure 3.5 - A.1 & 2) corresponds to the performance of the model when trained, tested, and validated using all locations (i.e., *prior information*). In the second strategy (Figure 3.5 - B.1 & 2), the *Validation* is a particular held-out location (*unseen locations*).

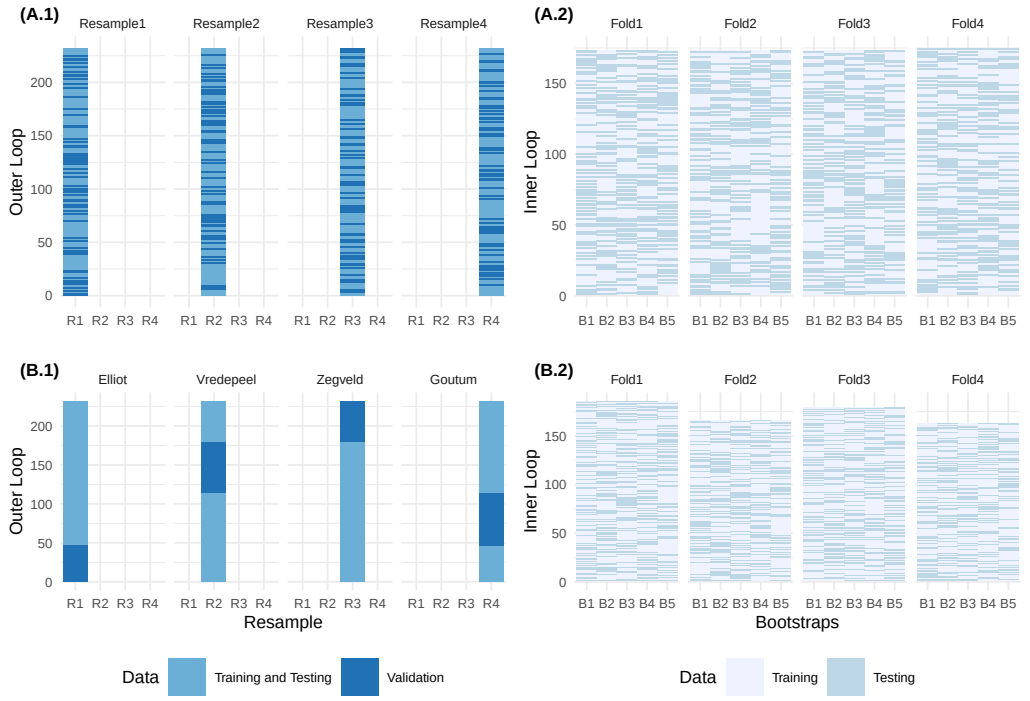



**Figure 3.5:** Validation Strategies. A.1 & 2: corresponds to a model validation strategy with *(a-priori)* knowledge, respectively presenting the Model Performance Assessment (A.1) and hyperparameter tuning (A.2). B.1 & 2 Model Performance Assessment in an unseen location. Hyper-parameters are tuned using bootstraps (B1 to B5) and validated against different Resample Folds (R1 to R4).

## 3.3 Results

### 3.3.1 Ground Truth Analysis

As presented in Table 3.2, the experimental design was appropriate for generating a large number of samples with a diverse range of biophysical and biochemical properties. The number ($n$) of observations acquired across all locations was equal to 266 (Elliot: 50, Goutum: 72, Vredepeel: 72, and Zegveld: 72). Following the removal of high biomass observations (i.e., sample-weight above 3500 kg DM/ha) and the outlier detection protocol, the final number of samples used for model development was reduced to 231. Most of the excluded data points (considered to be outliers or high leverage data points) belonged to Zegveld's data set (20 - Table 3.2). On closer inspection (i.e., digital photography), these detected outliers had a mixed botanical composition (white clover or *Trifolium repens* and perennial ryegrass), thus justifying its identification as outliers and removal of the data sets.

All locations displayed similar ranges for both %CP and biomass (Table 3.2), including both low and high values of %CP (e.g., between 10% and 27% %CP) and biomass (i.e.,

**Table 3.2:** Descriptive statistics of data biophysical and biochemical attributes per location:

| | Samples ($n$) | | $CP_m$ (kg CP.ha$^{-1}$) | | | % CP (% DM) | | | Biomass (kg DM.ha$^{-1}$) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Location | Initial | Final | Mean | Min | Max | Mean | Min | Max | Mean | Min | Max |
| Elliot | 50 | 46 | 223 | 43 | 578 | 20 | 12.3 | 27 | 1115 | 224 | 2986 |
| Goutum | 72 | 68 | 208 | 76 | 601 | 16 | 9.4 | 28 | 1380 | 526 | 3282 |
| Vredepeel | 72 | 65 | 172 | 70 | 479 | 18 | 9.6 | 26 | 1015 | 395 | 3223 |
| Zegveld | 72 | 52 | 283 | 70 | 664 | 18 | 11.5 | 24 | 1566 | 592 | 3420 |

between 500 and 3000 kg DM/ha). The sampling process at Elliot and Zegveld showed a uniform distribution across the $CP_m$ range, whereas the observations of Goutum and Vredepeel were skewed towards values below 200 kg CP/ha(Figure 3.6 - B).

Regarding comparability between locations, Bartlett's test ($p$-value <0.01) rejected the $H_0$ (null-hypothesis) under which $CP_m$ grouped by *locations* displays the same variance and, in context, were non-normally distributed. Given the non-normality, the KW test further indicated that there was a significant difference between distribution ($p$-value <0.01). Finally, the Dunn's test (using the Bonferroni adjustment method) indicated that while Vredepeel and Zegveld have different distributions, Elliot and Goutum were not significantly different ($\alpha$-level= 0.05) from any other location (Figure 3.6 - A).

*Pearson correlation* ($\rho$) between %CP and biomass was equal to -0.24 ($p$-value <0.01), indicating that both parameters were not collinear, a possible factor for spurious regression models. As expected, the correlation between $CP_m$ and biomass was both closely related and significant ($\rho = 0.91$, $p$-value <0.01). Lastly, Pearson correlation between %CP and $CP_m$ was equal to 0.13 ($p$-value = 0.05), indicating that higher values of $CP_m$ were not strongly associated with increases in %CP and were, borderline, not significant.

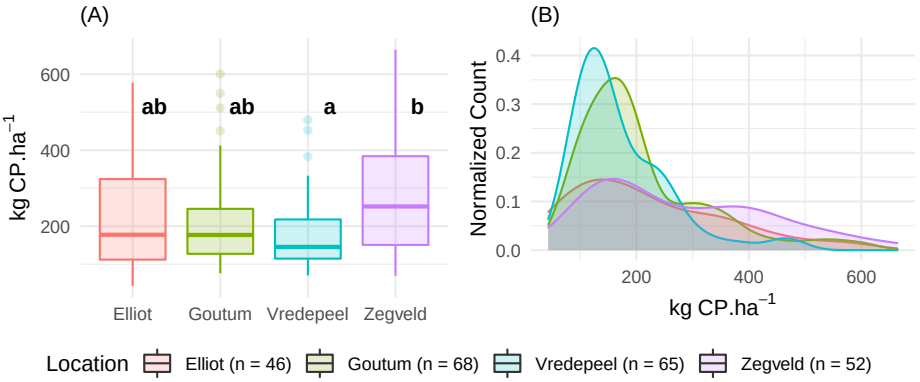



**Figure 3.6:** Boxplots (A) and density plots (B) of crude protein yield (kg CP/ha) grouped by different locations. Dunn's Multiple Comparisons ($\alpha$-level = 0.05) results presented next to each boxplot. Number of observations ($n$) per location within the bottom legend.

### 3.3.2 Spectral Analysis

The diverse range of biophysical and biochemical generated a wide range measurable of spectral responses, which was further enhanced by the pre-processing protocol (i.e., Savitzky-Golay smoothing filter and removal of noisy spectral areas, Figure 3.2).

An exploratory analysis was performed by grouping and averaging spectral measurements within each $CP_m$ quartiles (98, 151, 219, and 409 kg CP/ha). The relationship between $CP_m$ and reflectance levels showed to be consistent throughout the spectrum (Figure 3.8 - a to c). Specifically, in two distinct ranges (from 400–750 nm and 1475–2500 nm), a negative relationship between reflectance and $CP_m$ was observed. Conversely, from 750–1325 nm, $CP_m$ and reflectance levels were positively related. On the contrary, signs of saturation between the two largest quartiles were evident in two spectral ranges (400–560 nm and 1475–2500 nm). The largest separation between spectra was located within the NIR portion (750–1325 nm) of the spectra (Figure 3.8 - c). Both data sets (i.e., VIS-NIR and SWIR) employed in the GA feature selection processes, presented both direct and inverse relations between $CP_m$ and reflectance values. This allowed the regression algorithm to make use of these relationships without any explicit disadvantage when constraining the search-space to either VIS-NIR or SWIR.

When employing the same exploratory analysis to %CP (quartiles: 12.9, 16.4, 19.1, and 22.7%), such separation was not clearly nor linearly distinguishable (Figure 3.8 - d to f).

*Feature Selection (Band) Selection*

*Genetic Algorithm* - given the massive computational load, at most folds and spectral ranges, the algorithm was able to narrow down the optimal number of features within approximately the first 50 generations (Figure 3.7). In most folds, similarity (as measured by the Jaccard Similarity, Jaccard (1901)) between optimal individuals reached 90% in the $50^{th}$ generation. However, similarity between folds was not an issue and a comprehensive pool of features were selected for the final GA run.
Additionally, given the extensive sampling process of the training (i.e., $n = 1500$) and test (i.e., $n = 150$) sets - and assuming normality of residuals - 95% of residuals should fall within the average RMSE $\pm 1.96$ standard deviation ($\mu \pm 1.96\ \sigma$), allowing to infer whether predicting accuracies within and between spectral ranges were different (Welch $t$-test, *stats::t.test*).
Lastly, the initial number of bands available to the start of each GA process was equal to 70, 91, and 161 for VIS-NIR, SWIR, and the Full Spectrum, respectively. In all processes, the algorithm was successful in decreasing either the number of features or error metric, or both. To avoid excessive repetition and long numeric sequences, selected bands are provided as a *supplementary material*, although they are illustrated in Figure 3.8.

**Table 3.3:** Summary of Genetic Algorithm routine results for $CP_m$, modeled using different spectral ranges.

| | **Training** (RMSE - kg CP/ha) | | | | | | **Testing** (RMSE - kg CP/ha) | | | | | | **Features** ($n$) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\mu$ | Max | Min | $\sigma$ | Ini | Fin | $\mu$ | Max | Min | $\sigma$ | Ini | Fin | $\mu$ | Max | Min |
| VIS NIR | 45 | 50 | 43 | 0.97 | 47 | 45 | 49 | 63 | 35 | 7.2 | 50 | 49 | 11 | 24 | 8 |
| SWIR | 65 | 79 | 60 | 3.23 | 71 | 64 | 69 | 93 | 56 | 7.6 | 71 | 68 | 14 | 33 | 9 |
| Full | 48 | 55 | 45 | 1.54 | 50 | 47 | 50 | 80 | 37 | 8.7 | 53 | 50 | 20 | 52 | 13 |

*Note:* Confidence intervals ($\alpha = 0.95$) are expected to be within the mean ± 1.96 standard-deviation ($\mu \pm 1.96\sigma$).

### VIS-NIR feature selection protocol

**Crude Protein Yield ($CP_m$)** - The confidence interval ($\mu_{training} \pm 1.96\ \sigma$) for the training set was equal 45.1 ± 1.6 kg CP/ha (Table 3.3). Within the test set, the confidence interval ($\mu_{testing}$) lies within 49.3 ± 14.1 kg CP/ha while results ranged within 35 and 63 kg CP/ha. The average RMSE of test and train sets are statistically different (as per a two-sided $t$-test, $p$-value <0.01), which may signal over-fitting. However, the true differences ($\alpha - level = 0.01$) between averages is expected to be within -4.7 and -3.7 kg CP/ha, a small difference. As the test set results present a wider standard deviation (Figure 3.7), these results may indicate a difference in precision rather than accuracy.

In comparison, the RMSE values (test set) of the initial and final eleven generations are almost identical ($\mu_{initial} = 49.5$, and $\mu_{final} = 49.4$ kg CP/ha). Such difference was not statistically different ($p$-value = 0.89); however, the average number of features decreased from 14.8 to 9.7, which indicates that the GA process was able to effectively reduce the number of features while keeping an equivalent level of accuracy and precision (Figure 3.7).

**Crude Protein as Dry-Matter Fraction (%CP)** - The confidence interval ($\mu_{testing}$) was equal to 2.7 ± 0.7 %CP, while results ranged from 1.98 and 3.55 %CP (Table 3.4). When examining the performance of individuals of the initial and final eleven generations, a two-sided $t$-test indicates that there was a significant improvement in performance ($p$-value <0.01). The performance improvement is approximately equal to 7.1% ($\mu_{initial} = 2.87\%$ and $\mu_{final} = 2.66\%$) while there is a simultaneous decrease in number of features (13 to 11, respectively) (Table 3.4).

In summary, the VIS-NIR spectral range presented both best accuracy (smallest mean RMSE) and highest precision (narrowest standard-deviation) for $CP_m$ while also requiring the smallest number of bands. For %CP, as per Table 3.4, it ranked second both in accuracy and precision while requiring the minimal number of features. It should be noted that

**Table 3.4:** Summary of Genetic Algorithm routine results for %CP, modeled using different spectral ranges.

| | **Training** (RMSE - %DM) | | | | | | **Testing** (RMSE - %DM) | | | | | | **Features** ($n$) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\mu$ | Max | Min | $\sigma$ | Ini | Fin | $\mu$ | Max | Min | $\sigma$ | Ini | Fin | $\mu$ | Max | Min |
| VIS NIR | 2.6 | 3.2 | 2.4 | 0.09 | 2.8 | 2.6 | 2.7 | 3.5 | 2.0 | 0.38 | 2.9 | 2.7 | 12 | 21 | 8 |
| SWIR | 2.8 | 3.2 | 2.6 | 0.11 | 3.0 | 2.7 | 3.1 | 4.2 | 2.0 | 0.48 | 3.1 | 3.1 | 12 | 27 | 6 |
| Full | 2.4 | 2.8 | 2.3 | 0.06 | 2.6 | 2.4 | 2.6 | 3.3 | 1.9 | 0.31 | 2.6 | 2.6 | 20 | 41 | 12 |

*Note:* Confidence intervals ($\alpha = 0.95$) are expected to be within the mean ± 1.96 standard-deviation ($\mu \pm 1.96\sigma$)

such higher level of RMSE was due to the average RMSE of initial generations; which were significantly different from the final generations (two-sided $t$-test $p$-value <0.01).

In contrast, throughout the selection process, the VIS-NIR was consistently the spectral range with highest precision (narrowest standard-deviation), although its accuracy (average RMSE) was similar to the Full Spectrum (Figure 3.7). If the first 25 generations are set aside, the performance confidence interval for the VIS-NIR range was equal to 2.58 ± 0.13 %CP, best portraying its overall performance (Figure 3.7).

### SWIR feature selection protocol

**Crude Protein Yield ($\mathbf{CP_m}$)** - The average error for the training set ($\mu_{training}$) was equal to 65.3 ± 6.3 kg CP/ha, presenting a range of 60 and 78.7 kg CP/ha, the largest between the three spectral ranges. Accordingly, the test set presented the highest and widest RMSE values ($\mu_{testing}$ = 68.8 ± 14.9 kg CP/ha), with results in the range of 55.5 and 92.6 kg CP/ha (Figure 3.7).

A two-sided $t$-test indicated that initial and final eleven test sets ($\mu_{initial}$ = 71 and $\mu_{final}$ = 68 kg CP/ha) were significantly different ($p$-value <0.01), indicating that the GA routine was able to simultaneously decrease the average RMSE and number of bands (14.7 to 11.4, respectively). The initial number of features included in pool for the final GA run was equal to 47 out of 91 features, providing an ample subset for selection. The final number of features in the optimal individual was 11.

*Crude Protein as Dry-Matter Fraction (%CP)* - the confidence interval of the training-set ($\mu_{training}$) was equal to 2.8 ± 0.2, while the test set ($\mu_{testing}$) presented a higher and wider level: 3.1 ± 0.9 %CP (Figure 3.7). The final GA run pre-selected 34 of the 91 available features. During all runs, the minimum and maximum number of features selected within all generations ranged between 6 and 27 bands, a range only less wider than the Full Spectrum range (Table 3.4). Of all three spectral ranges, SWIR performed the worst, presenting the lowest accuracy and precision for both $CP_m$ and %CP.

### Full Spectra selection protocol

**Crude Protein Yield (CP m )** - The minimal RMSE (test set) was equal to 48.8 kg CP/ha, reached at the 132$^{th}$ generation. The initial generation of the GA process selected,

at each fold, an average of 32 bands (Figure 3.7), reaching the minimal number of variables (i.e., 13) and maximum (i.e. 52) at the $149^{\text{th}}$ and $1^{\text{st}}$ generation, respectively. The final run of selection uses 69 out of 161 features available and selected 19 bands as features for the optimal subset.

The range of RMSE found within the test set was 44.8 and 55.3 kg CP/ha with an average ($\mu_{training}$) of 47.5 and standard deviation ($\sigma$) of 1.3 kg CP/ha. Consequently, the confidence interval (95% range or $\mu_{training} \pm 1.96\sigma$) of observations was within 47.5 $\pm$ 2.6 kg CP/ha. In addition, for the test data set ($\mu_{test}$), the RMSE range (95% of the observations) was within 50.2 $\pm$ 16.6 kg CP/ha). The averages (i.e. training and testing) can be compared using a two-sample $t$-test ($\alpha$= 0.99), by which the means was statistically different ($p$-value <0.05), although the true difference between means were rather small (0.7 and 6.1 kg DM/ha at $\alpha$= 0.05).

When comparing the RMSE (test set) values of the initial generations (with more than 13 bands, $\mu_{initial}$ = 53.3 kg DM/ha) and those of the final generations (with 10 or less bands, $\mu_{final}$ = 49.9 kg CP/ha), the RMSE difference was statistically different ($p$-value = 0.01) and true-difference should lie ($\alpha$ = 0.05) within 0.69 and 6.05 kg CP/ha.

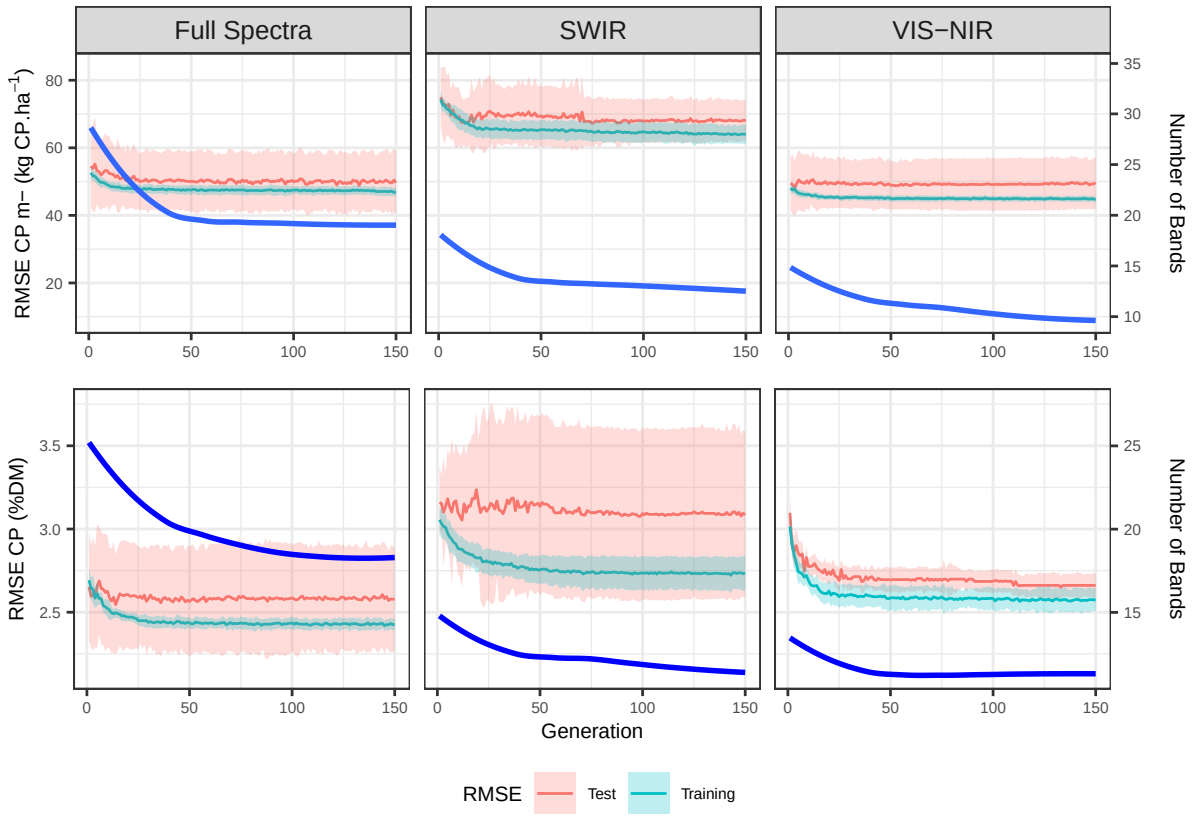



**Figure 3.7:** Performance during the Genetic Algorithm process. RMSE average and confidence interval ($\mu \pm 1\ \sigma$) primary axis. A locally estimated scatterplot smoothing (LOESS) of the number of bands (blue-color line) is presented in the secondary axis. Spectral ranges are presented in a column-wise fashion. RMSE for $CP_m$ and %CP are present in a row-wise fashion.

**Crude Protein as Fraction (%CP)** - models employing the full spectrum have consistently presented the largest number of bands. Initially, the optimal individual had on average 23 features. These individuals ranged from 12 and 41 features, while the smallest test RMSE was found using 25, although less than 20 features models had equivalent performance (Figure 3.7). Training and test range were equal to 2.3 and 2.8% and 1.9 and 3.3 %CP, respectively. For %CP estimations, on average, the FS range presented the smallest average RMSE from all spectral ranges, however its confidence interval was wider than VIS-NIR range (Figure 3.7). Confidence interval ($\mu_{testing} \pm 1.96\sigma$) for testing was within 2.6% $\pm$ 0.6% %CP (Table 3.4).

*Variable Importance*

The final feature sets selected by the GA process were also examined through variable importance proxies. The metrics for these proxies (Equation 3.1 and Equation 3.2) were considerably more elaborate than the internal maximization function used by the genetic

algorithm. Consequently, these methods were able to further indicate candidates for feature selection, as presented in Figure 3.8 (I to VI - right-side). Each spectral range was evaluated both by VIP (top) and sMC methods (bottom). The sMC thresholds (3.88) was the same across all models, given that samples size and consequently degrees of freedom were the same. For VIP all features below the threshold of 1 were considered not important.

*Crude Protein Yield* - Within the VIS-NIR (Figure 3.8 - I), the VIP method indicated that features below 700 nm were not important. The sMC indicated both 405 and 535 nm were not necessary for good model performance.
For the SWIR range (Figure 3.8 - II), the less important variables were within the 1600–1750 nm range, and the most important bands located in the NIR plateau (up to 1250 nm). However, throughout the bands of this spectral range almost no differentiation between each spectrum was found (Figure 3.8 - e).

In the Full Spectrum model, the features centered at 725, 775, 795, 1005, and 1045 nm were considered the most important according to the VIP method (Figure 3.8 - III). Likewise to the VIS-NIR range, the bands within the visible range 405, 535, 545, and 725 were identified as of lesser impact for the model's overall performance.

**Crude Protein as a Dry Matter Fraction (%DM)** - overall, the models for %CP perform poorly (Figure 3.10). Consequently, variable selection should be examined carefully or disregarded. As seen in Equation 3.2, such proxy (i.e. sMC) measures the difference in explained variance (already quite low in poor-performing models). Consequently, its interpretation was of limited-value.

Regarding the %CP of VIS-NIR, the features considered to be important were 425, 545, 685, 705, and 775 nm. In contrast to $CP_m$, all selected variables were mostly in the visible range and below 800 nm.
For the SWIR, no clear patterns can be identified as the VIP method selects 1155, 1205, 1305, 1525, 2135, and 2215 nm; while the sMC would not discard nor indicate a strong candidate for elimination.

Within the Full spectrum only 11 out of 25 have been considered important. Given that the VIP method is a translation of each feature weight across latent variables (Equation 3.1), the pattern exhibited in Figure 3.8 (VI), suggest that the model has incurred in overfitting, where contiguous variables present alternating high weights. Such condition (i.e., overfitting) was further evidentiated through the model performance analysis (Figure 3.10 - II and VI).

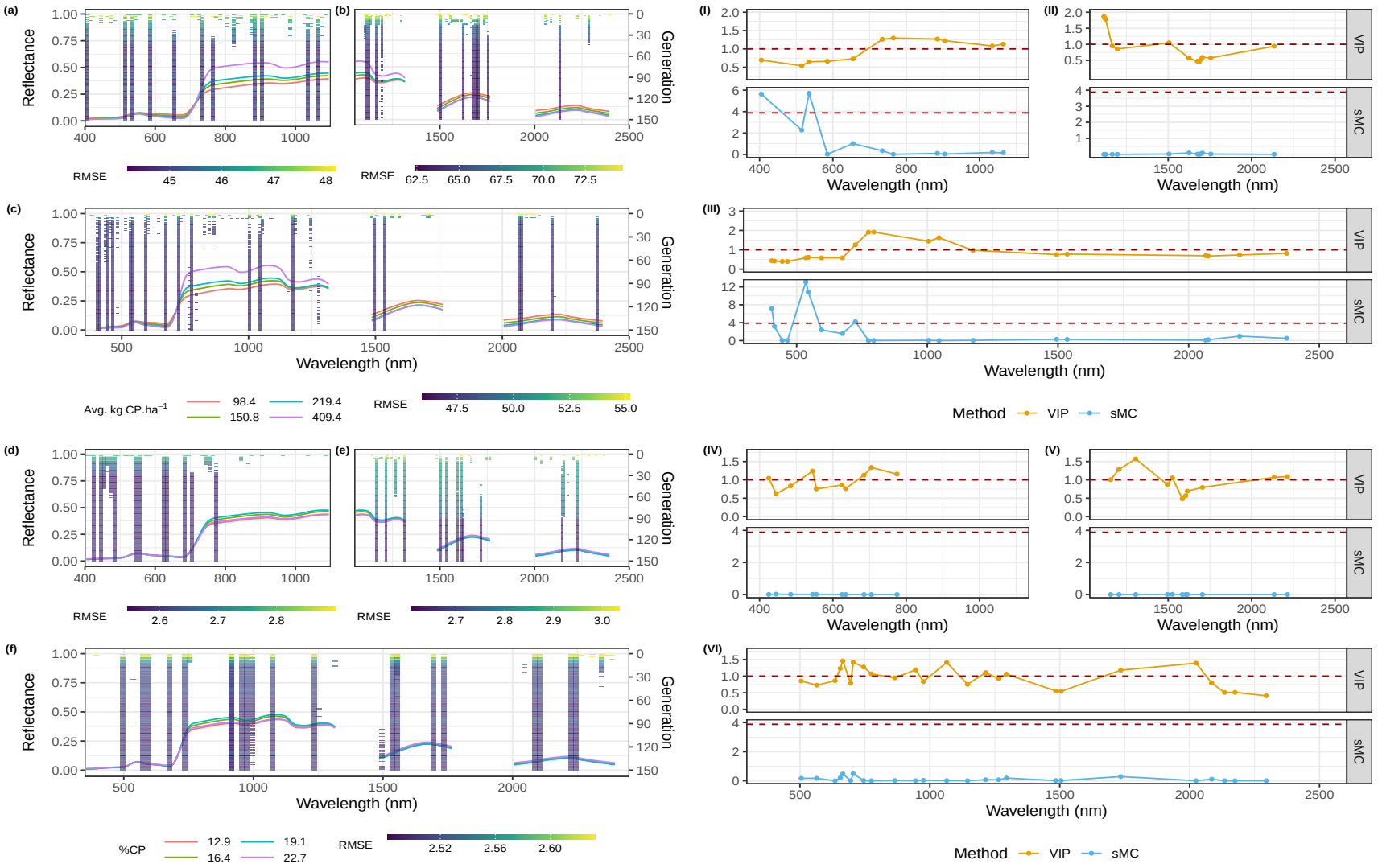



**Figure 3.8:** Feature (Band) Selection results of the genetic algorithm routine (left-side) and corresponding Variable Importance proxies (rigth-side). Upper-side and lower-side corresponding to $CP_m$ (a–c and I–III) and %CP (d - f and IV - VI), respectively

*Performance Assessment*

**Crude Protein Yield ($CP_m$) VIS-NIR** - The validation results for the models trained with prior knowledge was equal to 79.2 kg CP/ha (Figure 3.9 - I). Models presented a consistent behavior when validated in unseen locations, displaying an error equal to 91.6, 88.3, 90.5, 73.8 kg CP/ha for Elliot, Zegveld, Vredepeel, and Goutum, respectively. Overall, the error across sites was equal to 85.5 kg CP/ha (Figure 3.9 - IV).

For both **SWIR** and **Full Spectrum**, a large unequal residual dispersion was present up to 300–400 kg CP/ha (Figure 3.9 - SWIR and Full Spectrum). In contrast, the models based on the VIS-NIR present a higher level of homoscedasticity. Although the Full Spectrum presents a smaller RMSE for known locations, this may be seen as a random artifact of the seed value used for data partition given that both present the same average and standard deviation within testing data sets (Table 3.3).

As discussed, the VIP indicates that a possible overfitting for models based on the Full-Spectrum. Given such, the results for a more thorough testing (i.e., unseen locations), shows the algorithm was not able to keep the same level of accuracy. Its inaccuracies have a two-fold increase across validation strategies Figure 3.9 (III and VI).

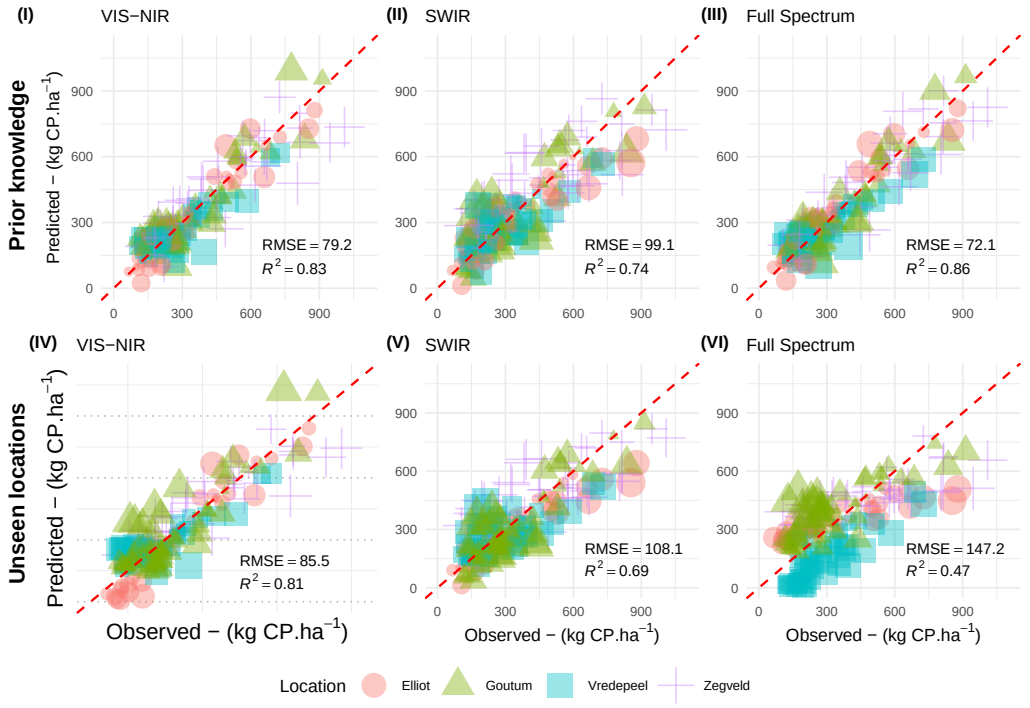



**Figure 3.9:** Predicted by Observed scatterplot using the bands selected through the GA band selection routine for $CP_m$ of each spectral range (column). Top-row corresponds to the PLSR algorithm trained in all locations. Bottom–row corresponds to PLSR algorithms trained withelding the test location.

**Crude Protein as Dry Matter Fraction (%CP)** - All models display a low level (i.e. $R^2$ <0.5) whenever assessed against unseen locations. Also, it is possible to notice that there is a significant clustering of predictions per location within (Figure 3.10 - V and VI). Between spectral ranges, for models trained with prior knowledge, results range from RMSE 2.6 to 2.7 %CP. For unseen locations, performances decreased to a range between 3.1 to 3.5 %CP (Figure 3.10). Similarly to $CP_m$, performances decreased when validated for unseen locations. From the three ranges, the VIS-NIR provided the smaller decrease in performance when assessed against unseen locations. The Full-Spectrum and SWIR had their coefficient of determination more than halved (Figure 3.10 - SWIR and Full Spectrum).

For all ranges, values above 25 %CP were poorly predicted.

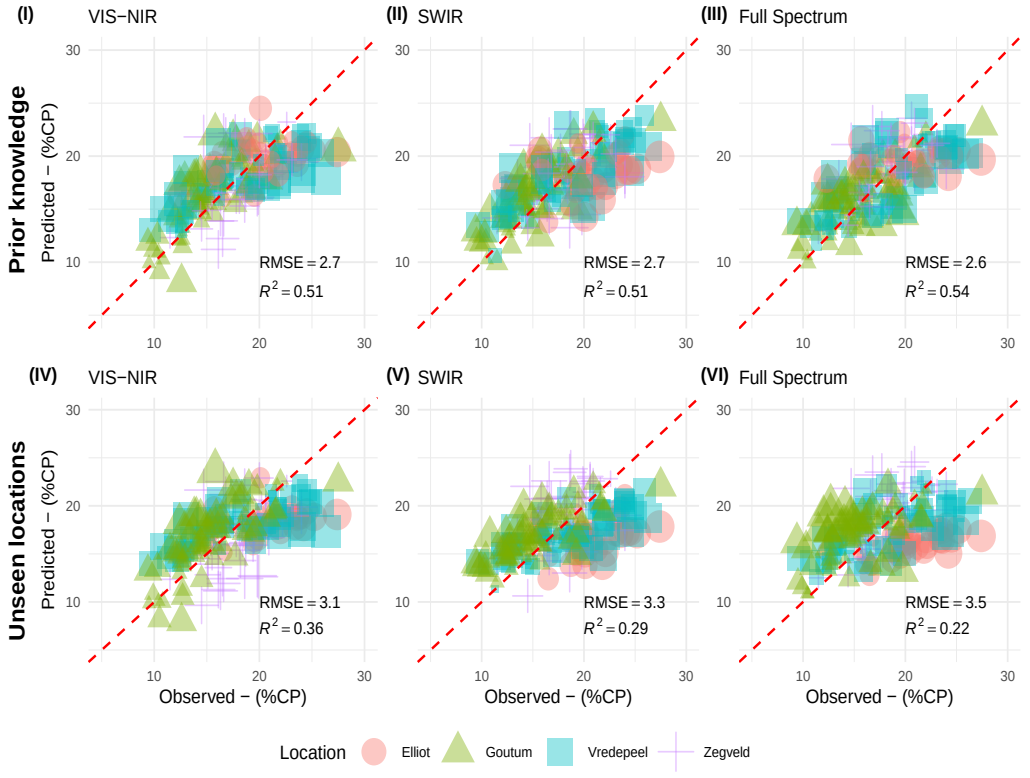



**Figure 3.10:** Predicted by Observed scatterplot using the bands selected through the GA band selection routine for %CP of each spectral range (column). Top-row corresponds to the PLSR algorithm trained in all locations. Bottom-row corresponds to PLSR algorithms trained withelding the test location.

## 3.4 Discussion

This study aims to establish an effective and affordable approach for a non-destructive, near-real-time, crude protein retrieval based solely on top of canopy reflectance. It has

successfully achieved its aims: firstly, the results showed that crude protein can be best explained, in field conditions, as a function of mass per area ($CP_m$) rather than as a fraction of biomass (%CP). Secondly, crude protein can be best estimated using the visible to near-infrared range of the spectrum (400–1100 nm), a range associated with low-cost *Si*-based sensors, instead of the shortwave infrared range (1100 - 2500 nm), as also indicated in Shorten et al. (2019) and Starks and Brown (2010). Thirdly, only a small number of bands (approximately eleven) are necessary for adequate model development. Despite a performance reduction, these (multispectral) models can be satisfactorily employed in unseen locations. Finally, these results indicate that a reasonably low-cost multispectral camera (or sensor) could - regardless of location - estimate perennial ryegrass $CP_m$ with an average error (i.e., RMSE) close to 85 kg CP/ha within biomass levels up to 3,500 kg DM/ha (Figure 3.9 - IV). It is reasonable to assume that these models could be further refined to a level close to 80 kg CP/ha RMSE (Figure 3.9 - I), provided that models are updated for new locations, similarly to NIRS laboratory methods.

This study has found equivalent results as Kawamura et al. (2010), however, employing a fraction of the number of bands proposed in that study. Kawamura et al. (2010) selected 167 narrow bands (FWHM = 1 nm) for $CP_m$ estimation, while this study indicates that eleven broad bands (FWHM = 10 nm) have an equivalent or superior performance (Figure 3.7). Our results are also in agreement with previous studies (Starks et al., 2016), both when indicating the best predicting spectral range (400–1100 nm) and the comparable prediction accuracy. Equally as important, for in-field conditions, the SWIR region has not consistently improved protein estimations, both when expressed as $CP_m$ or %CP (Figure 3.7 - SWIR and FS), not justifying its use for CP estimation. Such results extend the findings of Kattenborn et al. (2019), where the authors explicitly state that chlorophyll could be best estimate in its $\mu g.cm^{-2}$ form rather than in %DM basis.

Our results showed that %CP prediction error is within the range of 2.6–2.7% (Figure 3.10 - *"Prior knowledge" row*) for observations at known conditions, while high levels of %CP (i.e., above 22.7 %) are poorly predicted (Figure 3.10). In more rigorous assessments (i.e., unknown locations), inaccuracies are above 3 %CP. In contrast to Kawamura et al. (2010) and our results, Starks et al. (2006) report a higher achievable accuracy (i.e., around 1.5 %CP). Such a study was performed, however, for a different grass species (i.e., bermudagrass - *Cynodon dactylon*), under a smaller biochemical range (i.e., 4–14 %CP) and for a consistent canopy architecture (i.e., bermudagrass' leaves are oriented parallel to ground level - *planophile*).

In our regression models, most of residuals are within the largest %CP quartile (above 22.7 %CP). Similarly to Starks et al. (2006), the range below 15 %CP is well explained (Figure 3.10). Paradoxically, in well-managed pastures, perennial ryegrass' crude protein levels rarely drop below 18–20 %CP. However, a good spectral-separability can be found between the three first quartiles of observations (12.9, 16.4, and 19.1 %CP), allowing for

qualitative grading as per the Australian Fodder Industry Association (AFIA - Australian Fodder Industry Association Standards (2013)).

As reported in Kawamura et al. (2010), the feature selection routine (i.e., genetic algorithm) was shown to improve the model's final performance. In all cases, the procedure was able to simultaneously reduce the number of features and RMSE (Figure 3.7). In doing so, this routine highlighted important spectral regions, particularly, the range between the red-edge and the NIR-shoulder (Figure 3.8 - a and I).

The range of 700-800 nm has been consistently reported when estimating biomass and nitrogen content (Baret and Fourty, 1997; Mutanga and Skidmore, 2007), while being less prone to illumination effects (Sandmeier and Itten, 1999). Differently from other spectral regions, this region has a shift (i.e., the red-edge shift) towards wider wavelengths (Guyot and Baret, 1988), consistently linked with higher chlorophyll concentration and LAI, important and correlated features for CP. Such has been explored by Mutanga and Skidmore (2007) to estimate %N content with high accuracy for Buffel grass (*Cenchus ciliaris*) in a greenhouse experiment. The authors (Mutanga and Skidmore, 2007, p.41), expressly indicate the spectral ambiguity that could arise from different %CP and biomass levels as a significant limitation in their assessments.

These findings warrant the parsimonious selection of features rather than the common practice of arbitrary employment of all sensor's output. Although many regression algorithms should cope with multicollinearity or irrelevant variables, our results illustrate how a parsimonious model (Figure 3.9 - I and IV) is more generalizable than a model with a higher number of features (Figure 3.9 - III and VI). These results also indicate the need for judicious planning of validation strategies. While both the VIS-NIR and FS have a similar performance in an unconstrained $k$-fold cross-validation, the VIS-NIR model significantly outperforms the FS when using unseen locations as the validation set.

When examining the overall pattern of meaningful features for $CP_m$ (Figure 3.8 - I), it can be argued that these are convergent with the area of highest separation between spectra, found in the initial explanatory analysis (i.e., 750 - 1325 nm). As discussed by Kjeldahl and Bro (2010, p.559), the use of relevant variables (i.e., spectral range) is a crucial step in the development of efficient models. In a coherent fashion, most variables indicated to be important by the VIP are within the 725-1325 nm range (Figure 3.8 - a to c). Unfortunately, the arbitrary pre-selection of a spectral range is not straightforward as crude protein is present in a wide range of cellular constituents, and several overlapping absorption effects are present. Additionally, as shown in Figure 3.8 - I to VI, variable importance proxies are not always convergent and further selection could be addressed through different techniques based at experts' discretion and trade-off evaluation.

Nitrogen is a highly mobile plant-nutrient, present in many forms (Miller, 1935) and performing complex interactions within the plants. Given such complex interaction in a non-controlled environment, fundamental spectroscopy methods - such as those employed

in organic chemistry at the laboratory level - seem to be inappropriate in field-conditions. Consequently, one could hypothesize that the best strategy for practical purposes is to rely on correlation rather than causal effects. In other words, measurements are based not on molecular bonds (e.g. overtones), commonly associated with narrow band signals, but rather on the correlation of main drivers of absorption, broad band effects, such as plant pigments and cell-walls. Our best methods - although efficient - may be mapping (biological) correlations and not causal (e.g. overtones) indicators of protein (in its diverse forms) (Miller, 1935).

Our results show that the ability to estimate %CP at the paddock level solely from TOC reflectance is questionable or, at least, inferior than $CP_m$ estimation (Figure 3.7). Fundamentally, Beer's law directly relates light absorption as the product of concentration and optical thickness. To disregard a proxy of an optical path (e.g., biomass) will necessarily result in a underdetermined (or poorly determined) system. This ambiguity has been discussed in depth by different authors (Asner, 1998; Baret and Fourty, 1997; Curran, 1989; Guyot and Baret, 1988) and their findings are either in agreement or provide a strong background to support our results and limitations.

Although our experimental design does not aim to clarify these issues, it is reasonable to suggest that the measurement of CP as mass per area indirectly incorporate effects which are linked to optical path (e.g, LAI). Other sources of ambiguity, such as leaf angle distribution, may also be alleviated by the inclusion of biomass per area, given that despite not accurately determined, such factor presents a distribution of values per biomass range. Such underdetermination could better approached through the use of radiative transfer models. However, given the complex nature of necessary measurements, at this moment, such solutions seem inadequate for a farm-scenario.

It should be stated that, in laboratory conditions, the seminal NIRS work of Marten et al. (1989, Table 5) employed only 5-7 bands for %CP estimation without any spectral transformation, achieving a high level of accuracy (RMSEP = 0.5% and $R^2 = 0.99$). In field conditions, however, moisture content and canopy geometry will necessarily overlap and mask biochemical spectral features. In addition, the main advantage of hyperspectral data does not reside in the high number of bands, as contiguous bands are highly correlated, but in its application to derive spectral features such the continuum-removed features (e.g., band-depth, feature area) or resolving overlapping signals (Kokaly, 1999; Morimoto et al., 1999). However, although frequently and successfully employed in a laboratory setting, these techniques would require high radiometric accuracy in outdoor environments (and absence of the mentioned masking effects) rendering it unfeasible for precision agriculture purposes. Consequently, a custom-built multispectral systems - such as those proposed in this study - should be the most appropriate solution for an agricultural scenario.

*Limitations* - it is important to highlight that $CP_m$ and DM ($\rho = 0.91$) are highly linked; however it should not be understood as spurious relationship, given that biomass, LAI and

path length are components of TOC reflectance. However, these relationships should be best analyzed so causal relationships between biochemical components can be separated from biophysical effects. In essence, we advocate that %CP estimation (concentration) cannot be solved unless the TOC reflectance is corrected for biomass (path length) effects.

Another limitation of this study is the instrument's different point-spread-function for each of the three in-built sensors (VIS-NIR, SWIR 1 and SWIR2 ), as discussed in Mac Arthur and Robinson (2015). When not employing an optical scrambler, the response function within the instrument footprint is different for the internal sensors. Although most likely negligible, our experimental setup has not consciously controlled for these confounding factors or issues that could further enhance this problem (e.g., broken optical fibers or heterogeneity within the canopy). Imaging-spectroscopy and higher spatial resolution could clarify whether differences of %CP within the footprint may offset our method and enhance our understanding of the natural variability within the canopy. As noted by Curran (1989), the biochemical composition within a canopy is variable and how to best model its spatial variabilty *"requires the imaginative use of three-dimensional spatial statistics, giving the most weight to the uppermost layers in the canopy, rather than simple random sampling representation"*(Curran, 1989). Employing a hyperspectral imaging camera, Shorten et al. (2019) was able to provide a visualization of such while reporting differences in %N (equivalent to %CP) througout the canopy strata.

An additional limitation of the applicability of these findings and also a topic for future research are the scaling issues introduced by higher-altitude data acquisition (i.e., low-level flight) and different spatial resolution. Many of these issues have already been discussed by Wang et al. (2019) and Burkart et al. (2014), indicating that these can be overcome in an operational scenario. However, another important factor is the measurement error associated with imaging multispectral sensors (e.g. noise reduction, dark offset, vignette and row-gradient correction, and lens distortion), which are associated custom-made Multi-Camera 2D imaging system and have not been included in our method (nor it was the study aim). These limitations can be operationally resolved to a high level of radiometric accuracy, as demonstrated in Mamaghani and Salvaggio (2019).

From a sensing perspective, another important limitation is given by the selection of two bands beyond the 1000 nm (Figure 3.8 - I), concurring with the diminishing spectral sensitivity of *Si* after 1000 nm. Such constraint is trivial, however, given that spectral behavior from 800–1100 nm is highly correlated and driven by broad spectral features (e.g. cell-wall absorption feature).

## 3.5 Conclusion

This study has employed a modified feature selection technique, based on a genetic algorithm approach, to determine the best spectral region, the optimal and minimal

number of bands when estimating crude protein. Additionally, it has also contrasted how to best portray crude protein (either as dry matter fraction or weight per area basis) when utilizing top of canopy reflectance measurements for its estimation. Correspondingly, these tasks aimed to instruct the design of a new sensor (either point-measurement or imaging system).

Our results indicate that an affordable spectral-based sensor could estimate perennial ryegrass crude protein, in outdoor environments, using only eleven broad bands (10 nm bandwidth) within the visible to near-infrared range (400–1100 nm), provided that protein is expressed in weight per area. Additionally, models are transferable to new locations with a small decrease in performance (from 80 to 85 kg CP/ha RMSE). These results could lead to the development of sensors for autonomous deployment in an unmanned aerial or ground vehicle, providing end-users with essential information for best agronomic practices.

This study displays a rigorous comparison between how to best portray biochemical properties and model transferability under a robust theoretical and analytical framework. Finally, in light of future applications and technology adoption, it is necessary to re-evaluate how feed quality parameters are expressed whenever canopy-reflectance measurement is the ideal technique for data collection. Differently from benchtop near-infrared analysis, for field-conditions, biochemical parameters should be expressed as a conjugation of its concentration and optical path (e.g., weight per area, kg/ha) rather than the standard laboratory results (i.e. concentration, % dry matter).

To all effects, using estimates of crude protein as mass rather than a fraction of biomass does not limit the management of the herds' diet nor fertilization inputs. By expressing crude protein in its weight per area basis, the method and instruments presented can be used as a monitoring tool. Future research should focus on the translation and validation of our findings to a multispectral imaging sensor. Finally, different feed quality parameters could employ the method proposed in this study to determine achievable prediction accuracies and a minimal number of bands necessary for their retrieval.

# Chapter 4

# Retrieval of Hyperspectral Information from Multispectral Data for Perennial Ryegrass Biomass Estimation

This chapter is based on:

G. Togeiro de Alckmin, L. Kooistra, R. Rawnsley, S. de Bruin, and A. Lucieer (2020a). "Retrieval of Hyperspectral Information from Multispectral Data for Perennial Ryegrass Biomass Estimation". *Sensors* 20.24, 7192. DOI: 10.3390/s20247192. URL: https://www.mdpi.com/1424-8220/20/24/7192

*Supplementary materials to this chapter can be found in the online publication.*

# Abstract

The use of spectral data is seen as a fast and non-destructive method capable of monitoring pasture biomass. Although there is great potential in this technique, both end-users and sensor manufacturers are uncertain about the necessary sensor specifications and achievable accuracies in an operational scenario. This study presents a straightforward parametric method able to accurately retrieve the hyperspectral signature of perennial ryegrass (*Lolium perenne*) canopies from multispectral data collected within a two-year period in Australia and the Netherlands. The retrieved hyperspectral data was employed to generate optimal indices and continuum-removed spectral-features available in the scientific literature. For performance comparison, both these simulated features and a set of currently employed vegetation indices, derived from the original band values, were used as inputs in a random-forest algorithm and accuracies of both methods were compared. Our results have shown that both sets of features present similar accuracies (RMSE $\approx$ 490 and 620 kg DM/ha) when assessed in cross-validation and spatial cross-validation, respectively. These results suggest that, for pasture biomass retrieval solely from top-of-canopy reflectance (ranging from 550 to 790 nm), better performing methods do not rely on the use of hyperspectral or, yet, a larger number of bands than those already available in current sensors.

## 4.1 Introduction

*Precision agriculture (PA)* and *Remote Sensing (RS)* employ spectral observations to estimate biophysical and biochemical properties of vegetation canopies (Weiss et al., 2020). With the recent popularization of *Unmanned Aerial Systems (UASs)* for PA, farmers are now able to employ *Multispectral (MS)* cameras for spectral data acquisition, allowing for a high degree of flexibility in terms of data collection interval and custom-designed sensing systems (Aasen et al., 2018). Such systems can potentially be employed in precision pasture management as a monitoring tool (Michez et al., 2020), mapping key attributes such as biomass (Allen et al., 2011). This potential, however, is dependent in the optimal design of sensors (Dubbini et al., 2017), particularly on the number, position and width of spectral bands as well as its radiometric accuracy (Mamaghani and Salvaggio, 2019), aiming to provide high-quality predictors.

Currently, while constrained to MS data, PA practitioners rely on *Vegetation Indices (VIs)*, rather than individual band reflectances, as enhanced predictors of the canopy biophysical and biochemical properties (Rouse et al., 1973; Tucker, 1979). Overall, indices are summary measurements of broader and more complex systems, acting as indicators of an underlying phenomena. An index, in essence, summarizes a fraction of the spectral variability while tracking changes in the canopy (Anderegg et al., 2020) without necessarily resorting to the complete spectral information. In this sense, *VIs* are a valuable trade-off between a small number of bands and the full information available in a spectral signature (Hansen and Schjoerring, 2003).

Many indices, however, are prone to become irresponsive (i.e., to saturate) at high levels of biomass (e.g., *Leaf Area Index (LAI)* $\geq 4$) (Carlson and Ripley, 1997; Sellers, 1985), regardless of a continuous response in different parts of the spectrum (Gitelson, 2004). Furthermore, a large number of indices are available and proposed for different applications (Xue and Su, 2017). Consequently, this approach requires domain-knowledge for optimal VI selection (Silleos et al., 2006) and measurements at specific bands, requiring purposely-built MS sensors, which are usually either not commercially available or cost-prohibitive for commercial applications.

In a complementary way, *Hyperspectral (HS)* data has been shown to provide a higher level of accuracy when estimating different canopy properties (Mutanga and Skidmore, 2004b; Mutanga and Skidmore, 2007), particularly those which are functionally linked to narrowband absorption features, provided that the sensor spectral resolution is smaller than the observed spectral feature (Curran, 1994). As a drawback, the necessary sensor to collect such data is usually more complex and costly (Michez et al., 2020). Furthermore, broadband absorption features may overlap and mask the discrete and informative narrowband spectral signal (Ramoelo et al., 2011), impairing straightforward analysis and the added benefits of HS data (Berger et al., 2020). Lastly, hyperspectral measurements display a high level of

multicollinearity (Dormann et al., 2013), providing redundant information, particularly in contiguous bands (Esbensen and Geladi, 1990). Yet, while the individuals band reflectance *per se* are highly redundant, other spectral features such as rates of changes (i.e., derivatives), areas (i.e., integrals) and *Continuum-Removal (CR)* transformations may best capture, disentangle and enhance the predictive ability of hyperspectral data (McClure, 1993; Morimoto et al., 1999).

The constraints of VIs and HS are often presented as a dichotomy: while the first is seen as coarse and of limited efficiency (Baret and Guyot, 1991), the latter is presented as highly costly and complex yet able to map subtle changes in biochemical and biophysical properties (Berger et al., 2020). In many instances, VIs provide a partial understanding of the full spectral behavior, only directly displaying a fraction of the information contained in a spectral signature.

Within the spectral range of most commercial sensors (i.e. visible to near-infrared, 500–900 nm), the spectral signature of canopies is largely driven by broadband features (e.g., chlorophyll content and LAI) (Imran et al., 2020), corresponding to known response-functions (Guyot and Baret, 1988). It is, thus, reasonable to hypothesize that through a few sparse spectral measurements, the in-between bands can be interpolated and the underlying hyperspectral information can be retrieved, provided that adequate functions are chosen to describe the overall spectral behavior.

This *proposed* method would allow the reconstruction the spectral signature while taking advantage of the sparse nature of bands in commercially available MS cameras, and the generation of non-directly observed spectral bands and features. A similar approach was employed by Jongschaap and Booij (2004), who simulated the wavelength ($\lambda$) range from 660–810 nm in order to estimate the red-edge position, using a form of logistic function. More complex approaches, such as Bayesian (Gevaert et al., 2015) and multivariate methods (Zeng et al., 2017) have also been successfully employed to retrieve hyperspectral information from multispectral measurements.

Specifically for pasture biomass estimation, Mutanga and Skidmore (2004b) described the use of *CR* to derive vegetation indices while consequently reporting better performances than the usual normalized ratio indices. Also, using an exhaustive search procedure, several authors have proposed a red-edge normalized ratio index to estimate LAI and biomass (Mutanga and Skidmore, 2004a; Togeiro de Alckmin et al., 2021; Vescovo et al., 2012). These indices, however, cannot be directly extracted from commercially available MS sensors (either satellite or UAS-based), as the necessary bands are not measured. Yet, although these indices are not directly available from MS observations, accurately simulated spectra would allow their indirect estimation, and possibly lead to better estimations of biomass.

In summary, this study aims to (i) simulate hyperspectral data (550–790 nm) from multispectral measurements which can be recovered from a commercially available camera;

(ii) from the the simulated spectral signature, derive a set of synthetic high performance indices and high-performing spectral features found in the scientific literature (Mutanga and Skidmore, 2004b); and (iii) assess predictions of perennial ryegrass (*Lolium perenne*) biomass of the *proposed method* against a common subset of commercially employed vegetation indices, referred to as the *current method*.

The performance of these two methods will be further contrasted based on a cross-validation and a spatial cross-validation procedures (Meyer et al., 2019), granting insight on their performance in known and new locations. As a final analysis, for a deeper understanding of the performance analysis, indices and spectral features will be clustered through a dendrogram, providing an in-depth analysis of their redundancy level.

If successful, the *proposed method* would expand the number of VIs and spectral features generated from readily available MS cameras, matching more costly and complex HS sensors when applied to top-of-canopy reflectance measurements for pasture management. Finally, the accurate retrieval of HS information of vegetated canopies, purely based in MS data, presents an important argument to the optimal number of spectral bands for sensor design and whether hyperspectral data, and its derived features, provide value-added in the task of pasture biomass retrieval.

## 4.2 Methods

### 4.2.1 Experimental Layout

The experimental field design (concerning nitrogen levels and regrowth periods) was chosen to generate a wide range of biophysical and biochemicals properties while being representative of typical pasture management conditions of each location (Oenema et al., 2012; Rawnsley et al., 2014). The trial was performed at the Tasmanian Dairy Research Facility[1] (Australia) and at three research stations (Goutum, Vredepeel and Zegveld)[2] affiliated to Wageningen University & Research (the Netherlands).

In Australia, the experimental layout was an array of 30 rainfed perennial ryegrass plots (dimensions of 2.0 × 7.5 m, with 0.35 m border at each side of the plot's longitudinal axis), arranged as two rows by 15 columns. Within each plot, six spectra-biomass sample-pairs were collected. Plots were managed under five different *Nitrogen (N)* levels (0, 25, 50, 75, and 100 kg N/ha) and two regrowth intervals of approximately 15 and 30 days, with three pseudo-replicates of these combinations. Nitrogen levels were chosen based on previous research (Rawnsley et al., 2014).

[1] Elliot (41°4′57.7″S, 145°46′22.0″E)

[2] Goutum (53°10′38.8″N, 5°46′27.0″E), Vredepeel (51°32′57.9″N, 5°51′54.0″E), and Zegveld (52°8′32.9″N, 4°50′23.5″E)

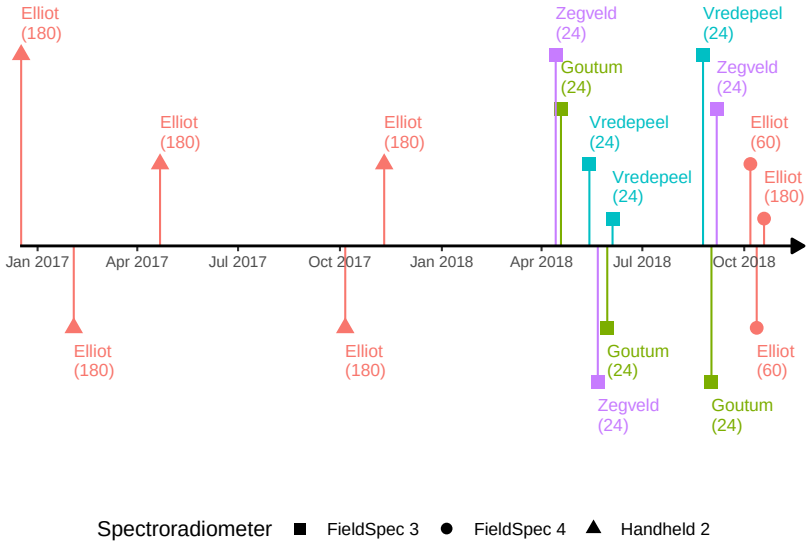



**Figure 4.1:** Data collection campaigns timeline. Each campaign is coded by location (color), number of observations (within brackets) and instrument (spectroradiometer) employed (shape). First date collection campaign took place in December 18$^{\text{th}}$ (2016) and the last, in November 24$^{\text{th}}$(2018)

Data collection campaigns took place on December 18$^{\text{th}}$ 2016, February 06$^{\text{th}}$, April 29$^{\text{th}}$, October 22$^{\text{nd}}$ and November 28$^{\text{th}}$ 2017, November 11$^{\text{th}}$, 17$^{\text{th}}$ and 24$^{\text{th}}$ 2018. In total, 1200 spectra-biomass sample-pairs were collected in Australia (Figure 4.1).

In the Netherlands, each experimental site contained 24 plots (dimensions of 20 $\times$ 20 m) and spectra-biomass sample pairs were taken at uniform canopy-cover and in the central area of each plot. The layout was a factorial combination of three N fertilization levels (0, 180, and 360 kg N/ha per year), four mowing intervals (in a cycle of four weeks, when then all plots were mown) and two pseudo-replicates of this combination. Each experimental setup was located within a different soil type (either clay, sandy or peat). Data collection spanned from May to October 2018. From the second half of July till September, plant growth was constrained due to a prolonged heatwave and drought. In total, 216 spectra-biomass sample-pairs were collected in the Netherlands (Figure 4.1).

### 4.2.2 Spectral Data

An identical data collection protocol was performed across sites and dates (Figure 4.2 - a to c). Spectral measurements were taken during clear-sky periods at around solar noon $\pm$2 hours, lasting for around 30 minutes in the Netherlands to one and half hours in Australia. To avoid a systematic illumination effect across plots, the order of spectral measurements of plots was randomized. An ASD FieldSpec®3 was employed in the Netherlands, a

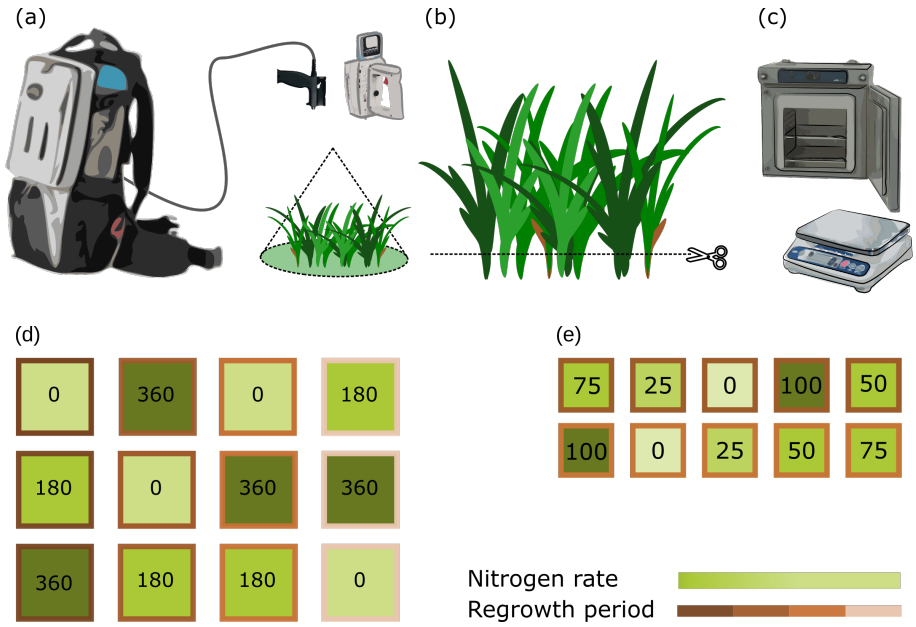



**Figure 4.2:** Data collection protocol (a - c) and plot layout (basic pseudo-replicate, d and e). Top: (a) Spectral Measurements, (b) Mechanical Defoliation and (c) Drying and Weighing. Bottom: basic Dutch (d) and Australian (e) plot layout. Borderline colors indicate regrowth period and hues of green (darker = higher rates) indicate nitrogen levels (rates are also indicated within each plot)

FieldSpec®Handheld 2 and a FieldSpec®4 (Malvern PanAnalytical - Colorado, Boulder, USA) were employed in Australia. The instruments had no attached foreoptics (i.e., bare fiber. Field of view: 25°) and were configured using the following setup (i.e., number of scans): 60 for white reference (Spectralon™. Labsphere - New Hampshire, North Sutton, USA), 60 for dark-current, 30 per measurement and five measurements of each target. Measurements were taken at nadir, from approximately one meter height, resulting in a circular footprint of approximately 0.44-meter diameter (Figure 4.2 - a). The instrument was recalibrated whenever the white reference measurement deviated from a straight line centered at one or a maximum time-limit of seven minutes between recalibrations was reached, whichever occurred first.

### 4.2.3 Reference Observations

The sensor's footprint was mechanically defoliated to a specific residual height (i.e., 50 mm) and stored in perforated plastic oven bags (Figure 4.2 - a and b). Such residual height corresponds to the best practice for perennial ryegrass management, reflecting a high level of light interception and common residual grazing height, consequently best portraying an operational scenario. The harvested material was immediately refrigerated and transported from the experimental sites to a forced-air oven, where it was dried for 48 hours at 65 °C and weighed (Figure 4.2 - c). In total, 1416 spectra-biomass sample-pairs were collected.

### 4.2.4 Data Analysis

Data was analyzed in R (version 4.0.2) (R Core Team and R Development Core Team, 2019) and for reproducibility purposes, data analysis R operations are introduced by the corresponding `package::function` format (typewriter typeface and accompanied by the double colon operator, i.e., the scope resolution operator).

### 4.2.5 Reference analysis

*Comparability between locations*: biomass measurements were tested for a series of *post hoc* tests (i.e., analysis of variance) to ensure that data collected between locations was comparable, testing whether average values ($\mu$) and variances ($\sigma^2$) were significantly different ($H_0$: $\mu_i = \mu_j$ at $\alpha = 0.05$). Initially, the Bartlett Test of Homogeneity of Variances (`stats::bartlett.test`) was performed, followed by Kruskal-Wallis (`stats::kruskal.test`) and Dunn's Multiple Comparisons (`FSA::dunnTest`), following the procedures described in Dinno (2015).

A Principal Component Analysis (PCA) was employed (`FactoMineR::PCA`) to decompose the variability of the MS data and visualize (`factoextra::fviz_pca_ind`) whether different locations grouped together in clusters, providing evidence whether or not spectral observations were comparable (i.e. not highly influenced by the soil background). For visualization purposes, ellipsis were drawn covering 95% of the observations of each location.

*Hyperspectral Simulation*

Spectral data was convolved (`hsdar::spectralResampling`) to the same specifications as a commercial multispectral sensor (Parrot Sequoia - Ile-de-France, France). Spectral resolution and the spectral response were extracted from its technical sheet using WebPlotDigitizer (Rohantgi, 2020). Four bands were generated: $B_{green}$ ($\lambda = 550 \pm 40$ nm), $B_{red}$ ($\lambda = 660 \pm 40$ nm), $B_{red\ edge}$ ($\lambda = 735 \pm 10$ nm), $B_{NIR}$ ($\lambda = 790 \pm 40$ nm).

*Piecewise equation* - The proposed method for hyperspectral simulation incorporated aspects of two previous studies (Jongschaap and Booij, 2004; Zeng et al., 2017). The first approach was presented in Zeng et al. (2017), where the authors use piecewise polynomials (i.e., splines) to interpolate the in-between spectra of a MS camera. Complementary, Jongschaap and Booij (2004) fitted a sigmoid function to the reflectance measurements of a CropScan™MSR87 (CropScan - Minnesota, Rochester, USA), showcasing the necessity of different functions types for different ranges of the spectra.

$$f(\lambda) = \begin{cases} \text{linear function} & \lambda < 685 \\ \text{logistic function} & \lambda > 685 \end{cases} \tag{4.1}$$

In the *proposed method*, hyperspectral simulation was based on an interpolation approach using a piecewise assembly of functions (Equation 4.1) to the convolved multispectral

dataset (Figure 4.3). In its first interval (i.e., sub-domain), the sub-function employed was a linear equation (Equation 4.2), in which the slope (i.e., $\beta$) was calculated using the $B_{green}$ and $B_{red}$ bands (550 and 660 nm, respectively). The second interval was based on a three parameter logistic function (i.e., K, $R_{680}$ or $B_{red}$, and C - Equation 4.3). Two of the parameters (i.e., K and $B_{red}$,) were the original band values, corresponding to the starting value (i.e., $B_{red}$) and asymptote (i.e., $B_{NIR}$) of the logistic function, respectively. The remaining and unknown parameter C (steepness rate) was found through a Nelder-Mead optimization approach (`stats::optim`) using three bands ($B_{red}$, $B_{red\ edge}$ and $B_{NIR}$). For each spectral observation, a set of parameters was computed and used to generate the simulated hyperspectral reflectance values (550–790 nm). No constraints were applied to ensure that the function was either continuous or differentiable. The piecewise function breakpoint was set to 685 nm (Equation 4.1 - usually the lowest reflectance value) and the logistic equation had its starting point at 660 nm (Equation 4.3).

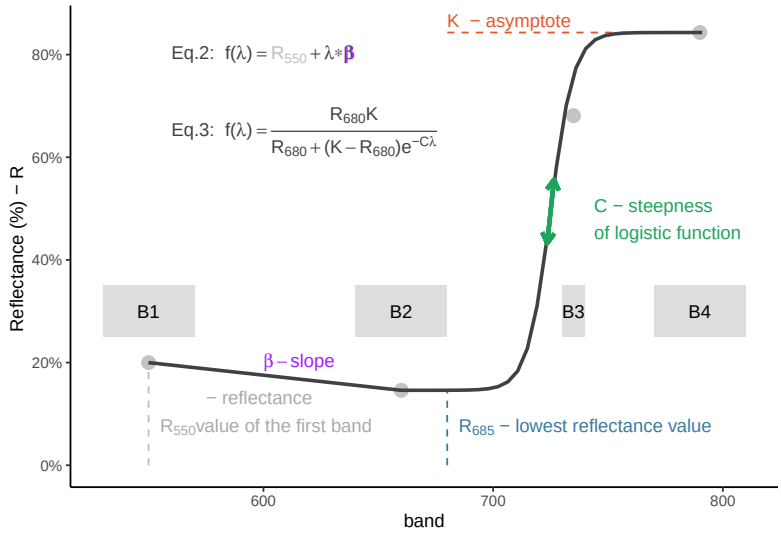



**Figure 4.3:** Hyperspectral simulation method. A piecewise equation Equation 4.1, using two different function: linear function Equation 4.2 and a logistic function Equation 4.3. Parameters were found through a constrained optimization procedure. Multispectral bands ($B_{1\text{-}4}$), corresponding to a common UAV multispectral sensor) displayed in grey shaded (width) area and grey circles (center)

$$f(\lambda) = R_{550} + \lambda \cdot \beta \qquad (4.2) \qquad f(\lambda) = \frac{R_{680}K}{R_{680} + (K - R_{680})e^{-C\lambda}} \qquad (4.3)$$

$\beta$ slope coefficient;

**R** Reflectance and $R_\lambda$ refers to the reflectance level at wavelength $\lambda$;

$\lambda$ wavelength (nm);

**K** logistic function asymptote level or $B_{NIR}$;

**e** the natural logarithm base (or Euler's number);

**C** the logistic growth rate.

*Spectral Data Retrieval from Multispectral Camera* - For the data collection campaigns of 17$^{th}$ and 24$^{th}$ November, 2018 (Figure 4.1 - Elliot), the multispectral camera was employed to collect imagery at low-level flight (i.e., below 130 m altitude) within ±1 hours of the solar-noon and under clear sky conditions. Flight duration was always under six minutes due to small area (0.1 ha) of the experimental area. The raw data was processed as per the Method "E" described in Poncet et al. (2019). The corresponding handheld sensor footprints were extracted from the imagery and their values were averaged. These averaged spectral responses were then employed as input in the hyperspectral simulation method, followed by a comparison with the handheld spectral data. For each date the number of observation pairs, between multispectral camera and handheld spectrometer, was equal to 180. Thus, in total, this analysis consisted of ($n$) 360 data points.

*Feature Engineering* - Features were generated as proposed in Kokaly (1999) (i.e., continuum-removal) and Mutanga and Skidmore (2004b) (i.e., band-depth indices). The authors reported that, of these band-depth indices, the best performing were based on *normalized band depth index* (NBDI), which is calculated by subtracting the maximum band depth ($D_c$) from the band depth (BD) and dividing it by their sum. These indices can be generated through the use of `hsdar::bdri` for Band Depth Indices and `hsdar::transformSpeclib` for Band Area (Lehnert et al., 2018).

Additional features properties (`hsdar::feature_properties`) were generated as well as an optimized normalized ratio vegetation index (NRI, $\lambda_1 = 745$ nm and $\lambda_2 = 755$ nm), which has consistently performed well in related studies (Mutanga and Skidmore, 2004a; Togeiro de Alckmin et al., 2021; Vescovo et al., 2012). From the CR properties, both $CR_{area}{}^3$, Max Depth Position$^3$ (i.e., wavelength position of the maximum value observed in the feature) were employed (Table 4.1 – Proposed Method).

Both $CR_{area}{}^3$, Max Depth Position$^3$ were employed through its cubic form to decrease the saturation effect. Finally, the logistic equation parameters found through the constrained optimization process (Equation 4.3) were also included in the subset (Table 4.1 – Feature Engineered).

The vegetation indices employed in the *current method* were: Normalized Difference Vegetation Index (NDVI), Green Normalized Difference Vegetation Index (GNDVI), Normalized Difference Red Edge Index (NDRE), Normalized Green Red Difference Index (NGRDI), Leaf Chlorophyll Index (LCI), and Structure Intensive Pigment Index 2 (SIPI2) as well as the original band values (Table 4.1 – Current Method).

**Table 4.1:** Table of predictors divided per Method

| Current Method | Proposed Method | Author (Reference) | Original Bands |
|---|---|---|---|
| GNDVI (Gitelson et al., 1996) | Optimized NRI $_{745\text{-}755}$ | Mutanga and Skidmore (2004a) | $B_{Green}$* |
| LCI (Zebarth et al., 2002) | $NDBI_{556 \text{ or } 744}$ | Mutanga and Skidmore (2004b) | $B_{Near\ Infrared}$ |
| NDRE (Barnes et al., 2000) | Max Depth Position$^3$ | Feature Engineered | $B_{Red}$ |
| NDRGI* (Jiang et al., 2019) | C and Beta | Feature Engineered | $B_{Red\text{-}Edge}$ |
| NDVI (Rouse et al., 1973) | $Band_{area}{}^3$ | Feature Engineered | |
| SIPI2* (Peñuelas and Filella, 1995) | $Band_{area}$ | Mutanga and Skidmore (2004b) | |

* Predictors marked with an asterisk were eliminated in the Filtering Process.

Given the large amount of possible features generated by the *proposed method* and the collinear nature of simulated spectra, the derived indices were chosen based on the best performing indices reported in Mutanga and Skidmore (2004b) (continuum-removal based indices), and a Normalized Ratio Index optimized for pasture biomass estimation (Mutanga and Skidmore, 2004a; Togeiro de Alckmin et al., 2021; Vescovo et al., 2012) .

Besides the band values and coefficients found within the piecewise regression, additional vegetation indices were: $Band_{area}$, Optimized $NRI_{745\text{-}755}$ , $Band_{area}{}^3$, Max Depth Position$^3$ (i.e., wavelength position of the maximum value observed in the feature) , $NDBI_{744}$, $NDBI_{556}$ (Table 4.1).

*Variable Filtering Selection* - For all features, including those developed in the *Feature Engineering* process, a minimal Pearson correlation-coefficient filter of $|0.2|$ was applied, following the methodology described in Alckmin et al. (2019). The remaining features were then employed in the modelling process (Table 4.1).

The formulas for the VIs can be found in the Appendix file. In Table 4.1, for the *current* methods, the references are next to each Index, and for the *proposed* method references are listed in the "*Author (Reference)*" column.

*Biomass Modeling*

Modelling (`tidyverse`) (Wickham et al., 2019) was performed employing the workflow described in Kuhn and Wickham (2020). Random-forest was chosen as the preferred regression algorithm for biomass retrieval as it was found to outperform other regression algorithms in a model performance study (Togeiro de Alckmin et al., 2021). The data was split in a 75%/25% ratio (training/testing and validation, using location as a stratification factor — Figure 4.4 - split). Explanatory values were centered and scaled. Hyperparameter tuning of the random-forest algorithm was found through a grid-search approach of two hyperparameters (i.e., number of trees and number of randomly chosen at each split) in a $k$-fold cross-validation approach (Figure 4.4 – Model Tuning). The chosen combination

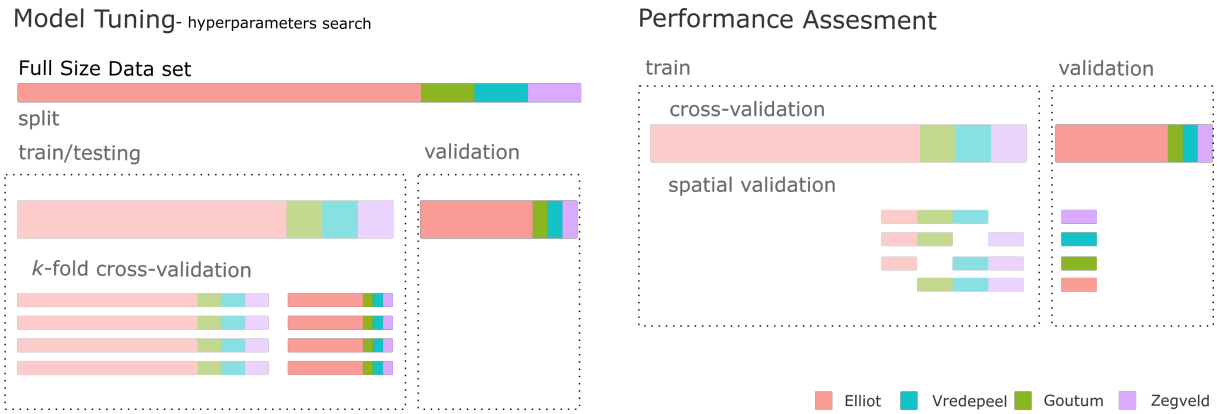



**Figure 4.4:** *Biomass Modeling* (left) and *Performance Validation* (right). Modeling: the original dataset was split in a train/test dataset (75%) and a validation dataset (25%). Tuning of hyperparameters was performed through a *k*-fold cross-validation approach. The best tuning (i.e. workflow) was then used in a *Performance Assessment* and validated through a (i) validation set or (ii) spatial validation strategy.

was within ten percent (`tune::select_by_pct_loss`) of the minimal Root Mean Square Error (RMSE) found within the search procedure.

The biomass modelling employed the same workflow, however, using two different subsets of VIs: (I) the use of original camera broadbands and VIs derived from these (referred to as *current method*); (II) using the continuum-removed features and the parameters estimate through the two functions of the piecewise equation (referred to as *proposed method*). For further comparison purposes, the *proposed method* was reproduced using the simulated and original hyperspectral data.

*Performance Validation*

**Model Performance** - Models performance was evaluated using two different strategies: (i) a cross-validation and (ii) and spatial cross-validation (Figure 4.4 – Performance Assessment). These two strategies differ in regards to the composition of training and validation sets. In the first (Figure 4.4 - (i) cross-validation), the algorithm was trained using the best combination of hyperparameters in the training dataset (75% of data). This model is then was assessed against a validation dataset equal to 25% of total observations (Figure 4.4 - Model Tuning).

In the spatial-validation, the original dataset was reduced so all locations have the same number of observations to prevent a class imbalance. Consequently, observations from Elliot were reduced from 1200 to 72. These observations were chosen at random while keeping the same biomass distribution as the original dataset. For the spatial validation, the algorithm was trained (using the same hyperparameter combination found in the model

tuning) and assessed four times, always keeping a different location as the validation set (Figure 4.4 - (ii) spatial validation).

**Feature Analysis** - Two unsupervised-learning visualization techniques were employed to provide insights about the dissimilarity amongst predictors and in relation to the predicted variable: the dendrogram and the heatmap, respectively. The clustering of the dendrogram was performed using the "Ward D2" hierarchical grouping technique (Ward, 1963). The heatmap organized predictors so as to match the clusters provided by the dendrogram. Additionally, the observations (i.e., predictor values) were sorted in a descending order, based on its biomass weight. The color pallete employed in the heatmap was scaled throughout predictors. The sorting (in descent order based on biomass values), alongside the scaled color pallete, makes explicit the behavior and the relationship between different predictors. This analysis allows insights on how dissimilar are the predictors (i.e.dendrogram), while also showing how predictors behave throughout the biomass range and if such behavior differs amongst them.

## 4.3 Results

### 4.3.1 Reference Observations.

**Biomass Measurements** – Although the number of samples ranges from 1200 (i.e., Elliot) to 72 (i.e., Goutum, Vreedepeel, and Zegveld), biomass ranges and distributions were comparable between all locations, with the exception of Vreedepeel. The Dunn test indicated that Vredepeel biomass distribution was not the same as other locations, as more than the 75% of samples present less 1500 kg DM/ha. All locations had 75% (i.e. three quartiles) of observations below the 3000 kgDM/ha (Figure 4.5 – a) .

**Spectral Data** – The inspection of PCA (Figure 4.5 – b) indicates that 98% of the MS variance could be decomposed in two principal components (i.e., Dim 1 and Dim 2). Observations were grouped per location (i.e., different colors and shapes) while different colored ellipses delimited 95% of observations of each location. These ellipsis were not separable (i.e., clusters), indicating that the grouping of spectra through its associated location has no distinguishable characteristics. In short, although each location has different attributes (e.g., soil type), by itself, the "location" factor has no influence in spectral-responses and cannot be used to point out differences in spectral response.

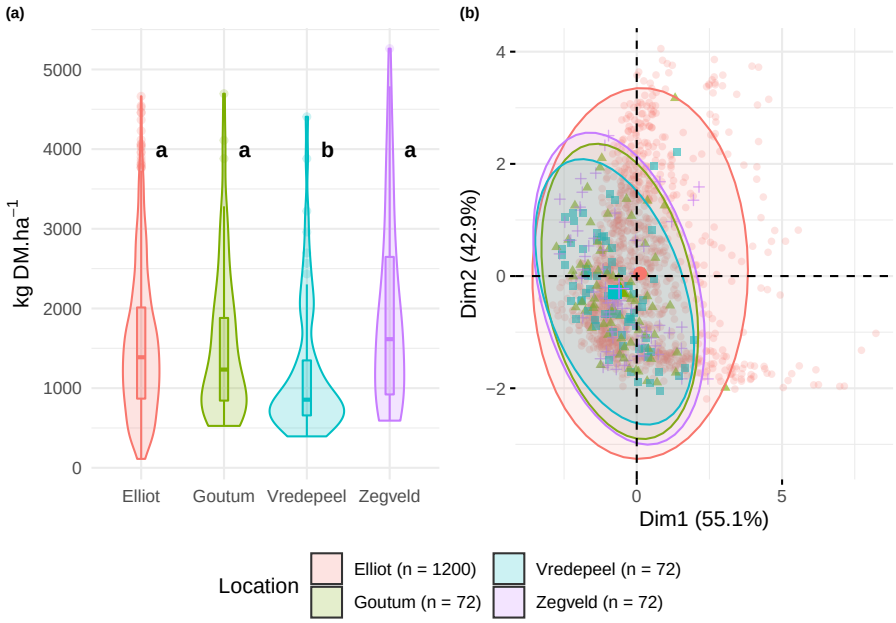



**Figure 4.5:** 1 - Boxplot and violin plot of observed sample biomass grouped by location. Dunn's Multiple Comparisons ($\alpha$-level = 0.05) results presented next to each boxplot. Number of observations ($n$) per location within the bottom legend. 2 - PCA of multispectral sample measurements. Ellipsis represent the centers of mass of each location and cover 95% of observations per group.

### 4.3.2 Data Analysis

*Hyperspectral Simulation*

Hyperspectral information could be retrieved with a high degree of correlation and small overall error (Figure 4.6 - a and b). Overall, simulated and observed spectral information showed a degree of correlation above 99% (Figure 4.6). However, some areas (e.g., red-edge and NIR plateau) consistently presented either over or underestimation of reflectance values, consequently indicating a higher level of RMSE. Yet, these average errors were usually below two percent (in reflectance values) (Figure 4.6 - a).

As the original reflectance values were retrieved with a high degree of accuracy, the CR could also be derived with satisfactory similarity to the original value (Figure 4.6 - b), allowing its use for the generation of predictors employed in the *proposed method*. As with reflectance, most of the simulated spectral range was highly correlated (i.e. above 97.5% correlation), although some areas (i.e. red-edge shift due to chlorophyll concentration) presented a higher level of error (i.e., RMSE).

Although minimal, an artifact was introduced at the breakpoint ($\lambda = 685\,\text{nm}$) due to the discontinuity of values output by the two sub-functions. Such artifact was more

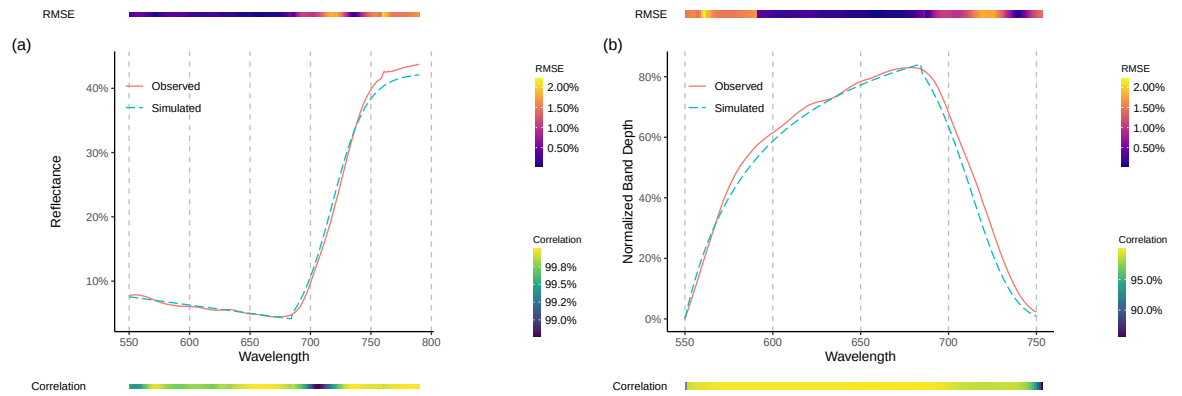



**Figure 4.6:** Observed and Simulated Reflectance (a) and Band Depth comparison (b). In each figure the corresponding average of all *Observed* and *Simulated* spectra are depicted. A top and bottom bar present both the RMSE (top) and correlation (bottom), per wavelength, between the *Observed* and *Simulated* spectra

pronounced in the continuum-removed features, seen as a sharp transition in its maximum point (Figure 4.6 - b).

*Spectral Data Retrieval from Multispectral Camera* - For both dates (17$^{\text{th}}$ and 24$^{\text{th}}$ November, 2018), there was a high level of agreement between the retrieved and observed spectral data, as presented in Figure 4.7. Overall, most bands presented a level of correlation above 80% and RMSE lower than 5% based on the ($n$) 180 observations per date. As suggested from Figure 4.6, the spectral regions corresponding to the red-edge inflection point ($\lambda \approx 730\,\text{nm}$) presented the lowest level of correlation, whereas the NIR region presented the highest level of RMSE. In addition, the artifact at the breakpoint ($\lambda = 685\,\text{nm}$) is noticeable. Between both dates, November 17$^{\text{th}}$ presented a better agreement between the retrieved and observed handheld spectral data (Figure 4.7 - right side).

*Biomass Modeling and Performance Validation*

*Feature Engineering and Filtering* – Within the *current* method, filtered VIs were: SIPI2 and NGRDI. Also, of the original MS bands, the $B_{\text{green}}$ was excluded from the modeling dataset. Regarding the *proposed* method, the filtering process did not exclude any of the initial predictor set (Table 4.1).

*Hyperparameter Tuning* – The grid-search of hyperparameters, both for the *current* and *proposed* methods, indicated that the less complex model (i.e., the one with a smaller number of trees and random variables per split) was within the ten percent RMSE model selection *tolerance* guideline (Figure 4.8). For all models, derived hyperparameters were 500 for number of trees and 2 for randomly selected predictors at each node. Performances within the train-test sets, displayed a marginal improvement of the *proposed* method

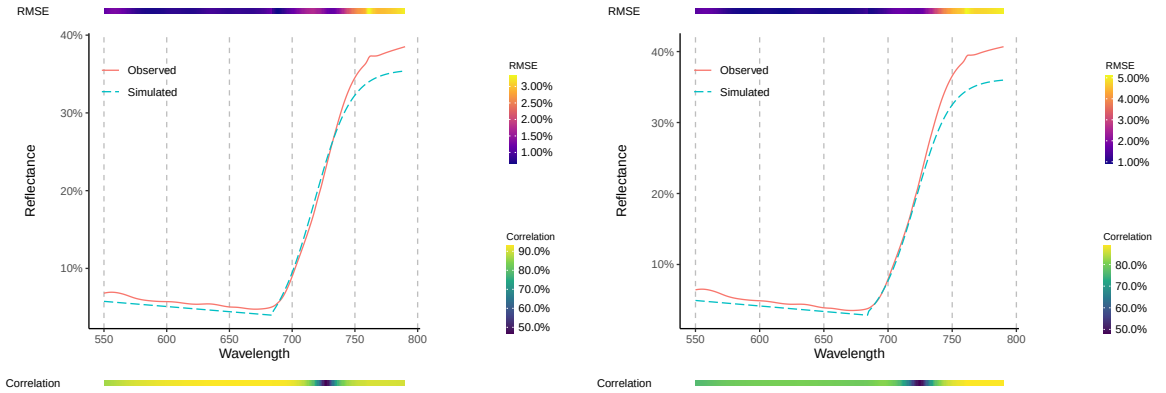



**Figure 4.7:** Observed Hyperspectral (solid red line) and Hyperspectral Retrieved (dashed blue line) from Multispectral Camera for the 17$^{\text{th}}$ (left-side) and 24$^{\text{th}}$ of November 2018 (right-side). In each figure the corresponding average of all *Observed* and *Retrieved* spectras are depicted ($n = 180$). A top and bottom bar present both the RMSE (top) and correlation (bottom), per wavelength, between the *Observed* and *Retrieved* spectra

(Figure 4.8 – a, b, d and e) in comparison to the *current* method (Figure 4.8 – d and f).

*Model Performance* – However, when applying the trained models to the validation set (Figure 4.9), in all cases, performance was equivalent. For the cross-validation approach, RMSE were close to 485 – 490 kg DM/ha. When employing the spatial cross-validation, performances decreased reaching values close to 610–620 kg DM/ha. Noticeably, in the spatial cross-validation (Figure 4.9 – d to f), there was a trend toward underestimation of biomass values for observations above 3500 kg DM/ha. This result, however, does not indicate a saturation of the prediction abilities of the model, as some observations in the 4000 kg DM/ha were accurately estimated.

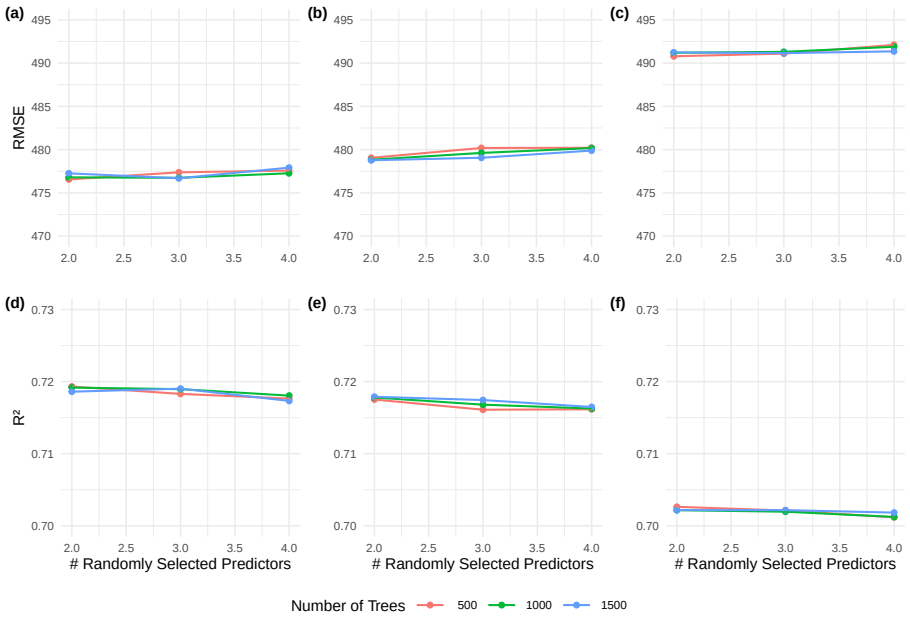



**Figure 4.8:** Grid-search hyperparameter tuning. Different coloured lines represent different number of trees and, the bottom-axis the number of randomly chosen predictor at each node. Sub-figures a and d are associated with training-testing values found for the *Proposed* method based on observed hyperspectral data, while sub-figure b and e correspond to the simulated hyperspectral data. Sub-figures c and f represent training-testing results for the *Current* method

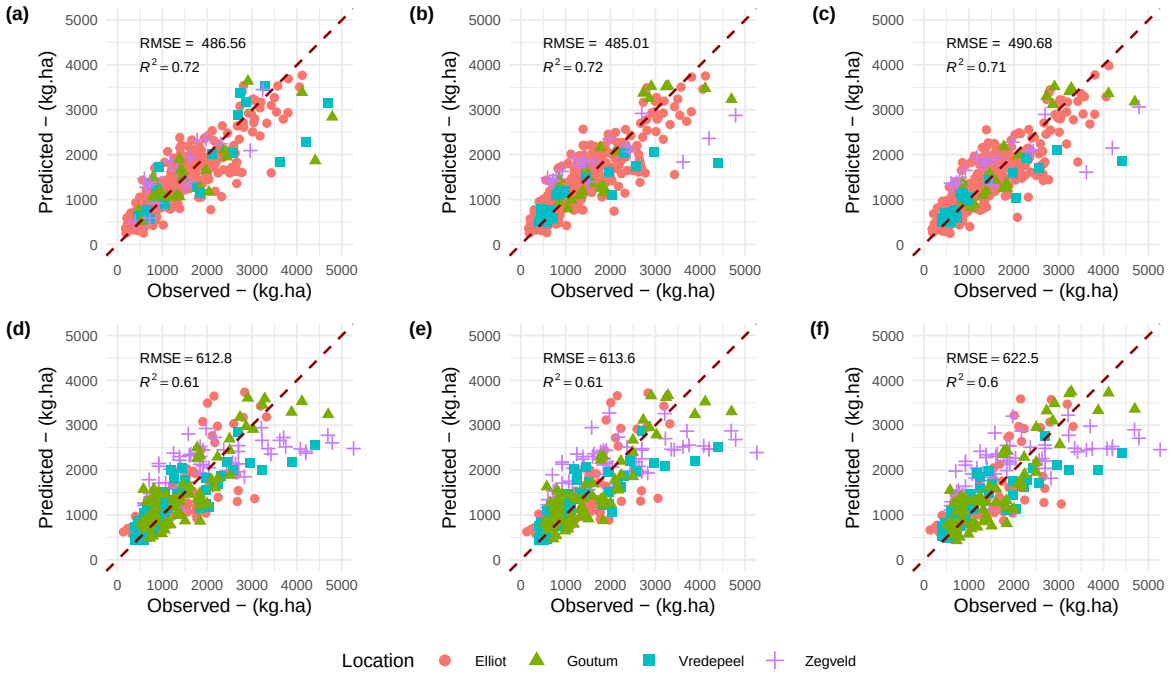



**Figure 4.9:** Predicted by Observed scatterplots. Validation strategy presented in a row-wise fashion. Accordingly, columns corresponds to the Methods employed. Top–row corresponds to cross-validation results. Bottom-row corresponds to spatial validation strategy. Methods: (a) – (d) Hyperspectral data, (b) – (e) *Proposed Method*, and (c) – (f) *Current method*

*Features Analysis*

Both Figure 4.10 and Figure 4.11 display the predictors clustered through a dendrogram (top-row), having at each end-leaf the corresponding correlation coefficient (linear relationship - $R^2$). Within the dendrogram, the higher the node-split, the most dissimilar are the variables between each branch and leaves, showing that the most dissimilar predictors were *Band Area* and *Red*. A threshold (light gray dash line) was established, and each branch is colored according to this division.

Within the heatmap, observations were sorted from top to bottom (row-wise) in a descending fashion, according to its associated biomass weight (Figure 4.10 and Figure 4.11 - bottom-right side). The value of each predictors was normalized (i.e., range equal to one) and colored accordingly (top-right side). On the right side of the figure, a scatterplot of the observations by corresponding biomass (kg DM/ha) is aligned (row-wise) with the corresponding predictors. As both sides of the figure are aligned, it is possible to observe the individual behavior of each observation in respect to the associated predictor values and biomass weight (Methods - Feature Analysis).

The dendrogram clustered the predictors in four main groups (different colors, according to the established threshold). The first group and largest group (i.e., purple branch) is formed

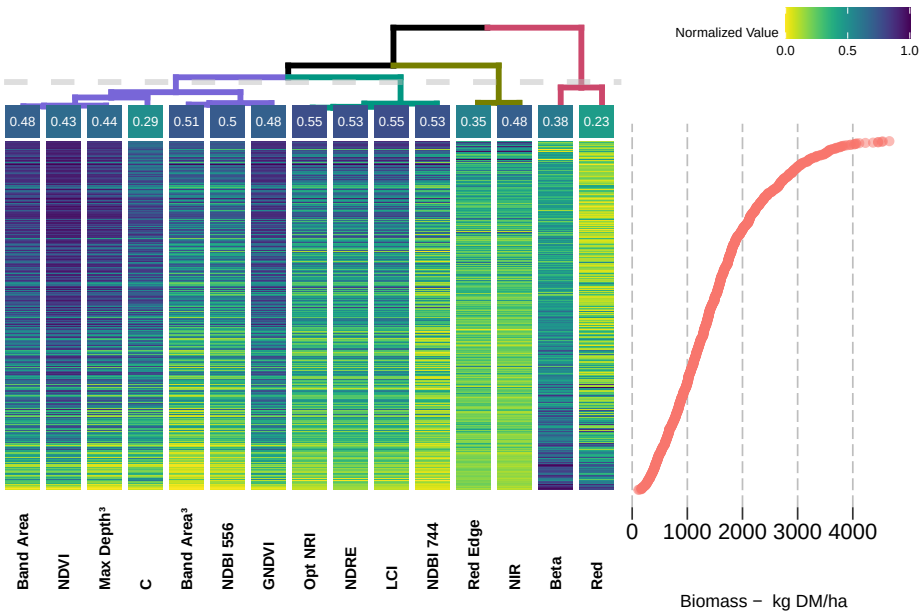



**Figure 4.10:** Dendrogram and clustering of features - Australia. Top left-side: dendrogram clustering of predictors, each end-leaf express the $R^2$ relationship with Biomass. Top right-side: color palette legend for the heatmap. Bottom left-side: heatmap of predictor variables, organized column-wise as per Ward-D2 criteria and row-wise in a descending order of biomass weights. Bottom right-side: scatterplot of the observations (i.e., samples) biomass weight, arranged in descending order

by Band Area (0.48), NDVI (0.43), Max Depth Position$^3$ (0.44), C (0.29), Band Area$^3$ (0.51), $NDBI_{556}$ (0.50) and GNDVI (0.48) – predictors and determination coefficients, respectively. In essence, this group clustered the predictors which were related to the VIS and NIR features and those related with CR features, and commonly associated with plant structural properties.

The second group (i.e., light-green branch) correspond to the predictors with the highest correlation with biomass, including the Optimized $NRI_{745\text{-}755}$(0.55), NDRE (0.53), Leaf Chlorophyll Index (LCI) (0.55), $NDBI_{744}$, which are commonly associated with plant greenness and different nitrogen concentrations. The following group (i.e., dark-green branch) is made of two of the original bands Red-Edge (0.35) and NIR (0.48). The final group (i.e., pink branch) presents the lowest correlation with biomass, and corresponds to the Beta (0.38) and Red (0.23). All four main branches present predictors from both *current* and *proposed* methods.

The proximity between leaves indicates a similar response-pattern between predictors and, consequently, biomass. Accordingly, distant branches correspond to less similar clusters.

The branches containing the original band values (i.e. dark-green and pink) are distant from VIs, displaying the two highest nodes.

Also, variables within the same branch display the same color-patterns, indicating its high level of collinearity. Finally, the response (color) pattern for the data collected in the Netherlands (Figure 4.11) is equivalent to the Australian data. In none of these heatmaps is possible to distinguish a particular behavior linked to Location. Through the visual inspection of the color pattern, it is possible to identify saturation effects (e.g. NDVI, GNDVI, $B_{red}$) of the predictors above 3500 kg DM/ha. In a opposite way, the indices clustered under the light-green branch (i.e.Optimized $NRI_{745\text{-}755}$, NDRE, LCI and $NDBI_{744}$) display no meaningful sign of saturation, reaching its maximum levels (i.e. darker-blue) at biomass levels above 3500 kg DM/ha.

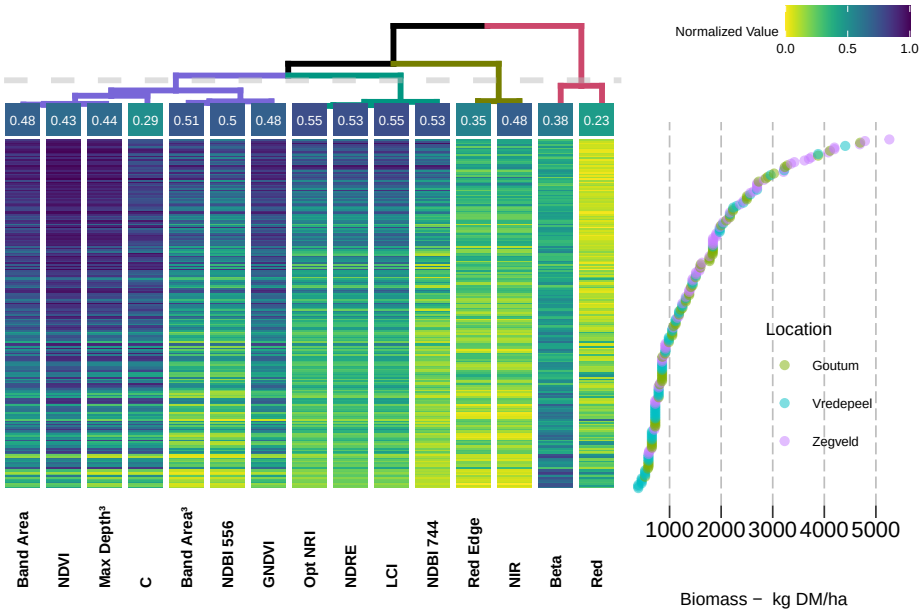



**Figure 4.11:** Dendrogram and clustering of features - Netherlands. Top left-side: dendrogram clustering of predictors, each end-leaf express the $R^2$ relationship with Biomass. Top right-side: color palette legend for the heatmap. Bottom left-side: heatmap of predictor variables, organized column-wise as per Ward-D2 criteria and row-wise in a descending order of biomass weights. Bottom right-side: scatterplot of the observations (i.e., samples) biomass weight, arranged in descending order

## 4.4 Discussion

Our results have shown that a satisfactory reconstruction of the hyperspectral signature of a vegetated canopy can be retrieved from multispectral measurements compatible with

a commercially available sensor (Figure 4.6 – 1). Such is to be expected, as the overall shape of the function (i.e., spectral response) is known and adequate functions were employed in the appropriate spectral ranges. Consequently, this study has also shown that a straightforward parametric equation can be employed for such purposes, without resorting to prior information as necessary for Bayesian or multivariate approaches (Results - Hyperspectral Simulation). When tested against a real multispectral camera, although with a smaller sample size (two different dates, $n = 360$), the hyperspectral retrieval method proved to be effective, reaching acceptable correlation and RMSE levels (above 80% and below 5%, respectively - Figure 4.7).

Furthermore, this study has shown that synthetic spectral features, such as continuum-removed features, can also be retrieved with a high degree of accuracy (Figure 4.6 – b). Moreover, models that employed either observed or simulated spectral data have performed equally well, achieving a RMSE of around 490 kg DM/ha, similar to values reported in Thomson et al. (2020) and Togeiro de Alckmin et al. (2021). The accuracy achieved reflects a long-term study, lasting close to two years, two different countries and with 1416 spectra-sample pairs.

Nevertheless, the results have shown that more elaborate VIs, such as those in the *proposed method* and originally introduced in Mutanga and Skidmore (2004a) and Mutanga and Skidmore (2004b), perform equally as well as those in the *current method*. Thus, it seems fair to state that "*while there is a long tradition of creating vegetation indices, there is also a tradition of proving these functionally equivalent*" (Perry and Lautenschlager, 1984). Such statement is corroborated through the analysis of dendrograms and heatmaps, where different VIs from both methods are clustered side-by-side within the same branches.

Both a narrowband optimal index (i.e., Optimized NRI) and a broadband index (i.e., NDRE), displayed almost identical correlation levels with biomass (Figure 4.10 and Figure 4.11). Also, both dendrogram and model tuning indicate that the number of predictors and model complexity could be decreased as no significant improvement was found when using more complex parameters (Figure 4.8) and many indices were shown to be redundant (Figure 4.10 and Figure 4.11).

Overall, results indicate that the maximum accuracy can be found using straightforward regression techniques, with a low risk of incurring in underfitting issues due to conservative hyperparameters employed (Figure 4.8). Consequently, it is reasonable to point that future research may explore less complex versions of the algorithm employed in this study: random forests. A number of different tree-based algorithms, may take advantage of non-linearity provided by this family of algorithms, while using a small number of predictors and less complex model structures. Such would may improve the overall understanding of the underlying phenomena, enhance its explainability to a broader audience and provide a higher scalability of methods to larger datasets while decreasing the necessary processing-power to deploy these models.

The range of error metrics found in the validation set are in line with previous research (Thomson et al., 2020; Togeiro de Alckmin et al., 2021). Based on the method and experimental design, an RMSE error equivalent to 460 – 495 kg DM/ha is equal to 7 – 7.5 g within the footprint. Although strict control was employed, such a small value may partially be due to different error sources, such as weighing scale precision, small imprecisions in cutting-height or incomplete harvesting. Such a systematic-error should be normally distributed throughout all-levels of biomass. However, larger residuals are found in high biomass levels (i.e., above 3000 kg DM/ha, indicating that other confounding factors (e.g. difference in scene support or ambiguous spectral response) may be decreasing achievable accuracies. A possible explanation lies in the well-known asymptotic nature of reflectance (Tucker, 1977a) by which a denser or thicker optical medium has decreasing reflectance response and, thus, further changes in the target do not equate to changes in reflectance levels (i.e. saturation). This heteroscedastic trend would be even more explicit if a larger percentile of biomass samples were above this saturation threshold. However, under operational scenario grazing or harvest of pasture are usually performed before this biomass range is reached.

From a biological perspective, two main broadband spectral features are acting within the spectral range (i.e., 550–790 nm) employed in this study: chlorophyll content and leaf area index. While, chlorophyll concentration is related positively to a minor deepening and a major widening of the absorption features, higher levels of LAI provide an asymptotic increase in reflectance level at the NIR plateau. These two phenomena explain the success in the retrieval of the hyperspectral data through a piecewise function and indicate that the ability to predict pasture biomass is linked with the ability to predict both phenomena regardless of the subset of vegetation indices employed to such end.

From a sensors design perspective, this study also suggested that the use of hyperspectral sensors in outdoor environment is of limited utility when not employed for measurements of narrowband absorption features. As an example of such, the simulated spectra cannot account for the fluorescence peak at around 760 nm or the red-edge position (Filella and Peñuelas, 1994) at 700–725 nm (Figure 4.6 – 1). Other areas of the spectra could be simulated with a high level of accuracy, thus not requiring a dedicated sensor.

As limitations of this study, although the handheld instruments employed are considered the benchmark for spectral measurements, commercial multispectral cameras can introduce instrumental error, due to poor design and radiometric processing, which may hamper the accurate retrieval of the spectral information.

Despite outside the scope of this study, the difference in between the simulated and camera-retrieved spectral data can be explained through differences in sensor-viewing geometry, illumination angle and radiometric processing of the multispectral data as well as differences in spatial support (i.e., differences related to the point-spread-function of the handheld instruments or regarding the footprint averaging of the multispectral imagery -

Mac Arthur et al. (2007)). Additionally, the radiometric consistency of the multispectral camera has been reported as sub-optimal (Franzini et al., 2019), possibly contributing to a fraction of the error. There was also a noticeable difference between dates, suggesting the need for further research on protocols for data collection of multispectral imagery at low-level flight.

As an additional limitation, soil spectral information has not been considered. Such, however, did not impacted the accurate retrieval of the original hyperspectral data. For most operational scenarios, grazing moments and pre-grazing biomass targets are reached after 95% light interception (Korte et al., 1988), thus target background, such as soils, has a negligible influence in the final spectral signature. The method proposed for spectral retrieval would be of limited utility for low-levels of biomass (e.g., below 500 kg DM/ha).

Conversely, it seems reasonable to hypothesize that, as reported in Tucker (1977a), at high levels of LAI, irradiance has little to no interaction with the lower strata of the canopy. Thus, for high levels of biomass, empirical relations between spectra and canopy properties have found its accuracy ceiling. Complex algorithm cannot improve performance due to an absent relationship between predictors and output. For this circumstances, accuracy improvements should come different sources of data or from different modelling strategies, such as time-series analysis.

Finally, this study successfully compared different validation strategies, both through a hold-out validation set and spatial-validation. These strategies differed in terms of accuracy, indicating that some variability is introduced in the relationship of spectra-biomass. Yet, there is no strong indication that location should be considered as a necessary predictor to improve accuracy, given the PCA analysis results. Furthermore, while there is a difference of around 120 kg DM/ha when employing spatial-validation, the same rational also shows that there are larger (i.e. 490 kg DM/ha, validation) sources of errors not specifically related to location.

## 4.5 Conclusion

This study has shown that hyperspectral data of vegetated canopies, in specific of perennial ryegrass, can be retrieved from multispectral measurements through the use of a straightforward piecewise parametric equation. As a limitation, however, discrete narrowband absorption features, such as fluorescence, could not be accurately retrieved. Vegetation indices and other spectral features derived from the simulated spectra have not improved model performance for biomass estimation when compared to indices already employed in current practice. Such may be due to the collinear nature of spectra and derived indices or transformations, introduced by the parametric equations.

The maximum achievable accuracy (i.e. RMSE) was equal to approximately 490 kg DM/ha. This error-level may be used as a benchmark for spectra- based models when employing top-of-the-canopy reflectance ($\lambda$ = 550–790 nm). Such limitation may be due to the asymptotic response-nature of reflectance and high levels of biomass (LAI > 4), and to uncontrolled sources of error (such as scene support) not accounted for in this study.

Finally, the results indicate that, in outdoor conditions, a multispectral sensor can perform equally as well as a hyperspectral sensor when aiming to retrieve pasture biomass estimations and the spectral range employed is limited to 550–790 nm. In this range, the main reflectance drivers are chlorophyll absorption and cell-wall reflection, both broadband phenomena and measurable through a multispectral sensor.

Future research should focus on additional uncorrelated predictors not correlated with spectral response, such as weather data, different modelling strategies, such as time-series or forecast models, and testing whether less-complex machine learning algorithm could provide the same level of error, while decreasing processing power requirements for model deployment.

# Chapter 5

# Perennial Ryegrass Biomass Retrieval Through Multispectral UAV data

This chapter is based on:

G. Togeiro de Alckmin, A. Lucieer, R. Rawnsley, L. Kooistra. "Perennial Ryegrass Biomass Retrieval Through Multispectral UAV data". Submitted 19/Jan/2021.

# Abstract

Frequent biomass measurement is a key-activity for optimal perennial ryegrass (*Lolium perenne*) management. Due to the necessary frequency and current trend of larger dairy farms, such activity is perceived as an operational bottleneck. Substantial effort is directed to the development of accurate and automated technological solutions. The recent popularization of unmanned aerial vehicles (UAVs) combined with multispectral cameras should allow for an optimal observational system able to deploy machine learning algorithms for near real-time biomass dry-matter (DM) mapping. For successful operation, these systems should deliver radiometrically accurate orthomosaics and robust models able to generalize across different periods. Ideally, such pipelines should require minimum processing power and allow for fast deployment. This study has established a two-year experiment, comparing reflectance measurements between a handheld spectrometer and a commercial multispectral UAV camera. Additionally, different algorithms based on regression-tree architecture were contrasted regarding accuracy, speed, and model size. Model performances were validated, providing error-metrics for baseline accuracy and temporal validation. The results have shown that the standard procedure for multispectral imagery radiometric calibration is sub-optimal, requiring further post-processing and presenting low correlation with handheld measurements across spectral bands and dates. Nevertheless, after post-calibration, the use of spectral imagery has presented better baseline error than the point-based sensors, respectively displaying an average of 397.3 and 464.2 kg DM/ha when employed alongside the best performing algorithm (Cubist). When trained and validated across different years, model performance was largely reduced and deemed unfit for operational purposes. The Cubist/M5 family of algorithms have exhibited advantageous characteristics such as compact model structure, allowing for a higher level of model interpretability, while displaying a smaller size and faster deployment than the Random Forest, Boosted and Bagged Regression Trees algorithms.

## 5.1 Introduction

Efficient pasture production and utilization are often the most critical components in a dairy operation (Wilkinson et al., 2020), directly impacting the overall profitability and carbon footprint associated with dairy systems (Lorenz et al., 2019). Moreover, the recent widespread adoption of *Unmanned Aerial Vehicle (UAV)* as a precision agriculture tool offers unprecedented opportunities for pasture biomass assessment (Michez et al., 2019). Such activity, which has historically been a bottleneck on forage-based dairy systems (Edirisinghe et al., 2012), can now conveniently capitalize on *Remote Sensing (RS)* techniques and semi-autonomous platforms for intra-paddock data collection and biomass assessment.

In the past decades, substantial progress has been achieved in determining key biophysical attributes of pastures through the analysis of spectral data, particularly through the use of field spectrometers (Kawamura et al., 2009; Mutanga and Skidmore, 2004a). Although demonstrating the potential of *in-situ* reflectance analysis, handheld instruments do not address the need for automated data collection necessary for an operational farm scenario. This challenge, however, can be surpassed through the use spectral imaging systems, such as *Multi-Camera 2D Imagers*, and UAV which are becoming ubiquitous in precision agriculture practices (Aasen et al., 2018).

Ideally UAVs should fulfill an observational gap, capable of providing the necessary spatial, spectral and temporal resolution for farm operations, translating handheld point-based methods to farm-scale aerial mapping. In terms of physical measurements (i.e., reflectance), the difference in data acquisition scale (i.e., from ground to low-level flight) has been shown negligible, yet subject to sensor quality and design as well as environmental illumination conditions (Von Bueren et al., 2015). Such equivalence between spectral data collected at canopy and low-level flight (i.e., below 120 m height) indicates the transferability of methods, providing an effective tool for measurement and monitoring of pasture attributes.

Currently, commercially available UAV imaging systems are limited to a small number of spectral bands and often rely on the use of *VIs* as enhanced predictors. Many of these indices, however, are prone to saturation, displaying an asymptotic response to high levels of biomass (Tucker, 1977a), limiting model performance due to heteroscedasticity (Muñoz et al., 2010). Yet, saturation and poor model fit have been shown to be reduced through either using optimized or a small subset of VIs, particularly when employing non-parametric regression models (Guerini Filho et al., 2020; Mutanga and Skidmore, 2004a; Togeiro de Alckmin et al., 2021), further improving the potential and competitive advantage of UAVs multispectral systems.

Widespread adoption of these tools is limited due to data size and processing power requirements, leading to onerous scalability to large areas in a fully automated pipeline. These constraints require algorithms which can be quickly deployed both for the generation

of orthomosaics to end-products (e.g., biomass maps) either in a cloud-based system or on a personal computer (Maes and Steppe, 2019).

In recent years, the generation of georectified mosaics has been largely overcome (Turner et al., 2012; Zhao et al., 2019) through the use of *Structure from Motion (SFM)* techniques, and is currently commercially offered. However, the processing pipeline for radiometrically corrected orthomosaics is under active research and improvement (Fallet and Domenzain, 2018; Gilliot et al., 2018; Suomalainen et al., 2018).

Data quality is subject to image processing from raw data to radiometrically corrected orthomosaics (Poncet et al., 2019) and systematic errors introduced by sensor design (Barker et al., 2020; Mamaghani and Salvaggio, 2019). Despite a substantial demand, end-user products such as quantitative biophysical maps, which are dependent on radiometric accuracy, are not (yet) in place. Inevitably, imaging systems must provide accurate radiometric data for reliable deployment of predictive models.

Physical models (i.e., radiative transfer models) have successfully described canopy spectral responses through the use of biophysical inputs (e.g., chlorophyll concentration, leaf area index). These biophysical inputs are also employed in pasture sciences as main factors for adequate levels of fertilization, photosynthetic activity or biomass accumulation, which are key performance indicators for optimal management. However, these drivers of spectral response are not easily measured in the field.

The use of interpretable models should allow end-users to have a real-time assessment of a key management parameter (e.g., biomass or nitrogen concentration) while simultaneously indicating the drivers for that estimation (e.g. chlorophyll concentration or leaf-area index) through links between spectral predictors and biophysical inputs (e.g., VI vs. canopy structure). Additionally, such type of information may be employed towards model diagnosis, establishing under which conditions (e.g. LAI $\geq$ 10) models tend to under perform, and in which conditions specific spectral predictors are of greater importance.

Meanwhile, there is usually a trade-off between (i) accuracy, (ii) model complexity and (iii) interpretability (Carvalho et al., 2019). Rule-based models, such as decision trees, are an interesting compromise between these three factors due to its interpretability and accuracy provided by its non-parametric nature. The fit between the canopy spectral-response (in the visible to near-infrared range) and biomass should not display a complex pattern, thus not requiring black-box complex models for optimal fit between predictors and explained variable. Non-complex models, such as shallow decision-trees should allow for interpretability and linkage between specific VIs and biophysical inputs.

Furthermore, the prediction performances of machine learning models are typically communicated via $k$-fold cross-validation estimates, possibly providing an optimistic assessment (Roberts et al., 2017). A more far-reaching evaluation of performance explores

spatial and temporal assessments (Meyer et al., 2018; Togeiro de Alckmin et al., 2020a), providing insights to end-users about performance under different circumstances.

The objective of this study is to (i) assess the correlation between reflectance measurements of a commercially available UAV multispectral camera at low-level flight and ground-based measurements of perennial ryegrass *Lolium perenne*, (ii) assess the difference in accuracy, size and speed when employing gradually more complex tree-based regression algorithms and (iii) validate model performance through different strategies: (a) $k$-fold cross-validation, and (b) validation in different years. The findings from this study provide a foundation for the assessment of UAV remote sensing for precision pasture management applications as well as a strong framework for biomass monitoring.

## 5.2 Methods

### 5.2.1 Data Collection

The experimental field trial was undertaken at the Tasmanian Dairy Research Facility in Elliot (TAS, Australia — 41°45′7.3″ S, 145°46′21.8″ E). The experimental layout was an array of 30 rainfed perennial ryegrass plots (dimensions of 2.0 x 7.5 m, with 0.35 m border at each side of the plot's longitudinal axis), arranged as two rows by 15 columns.

Plots were grouped in three main blocks (10 plots per block). For the 2017 data collection campaign, each block was split in different growth intervals: long and short or approximately 30 and 15 days, respectively. For the 2018 data collection campaign, each block had a single growth period. All plots were mown and fertilized in October 21$^{st}$, 2018. In the next two subsequent data collection campaigns, a single block was harvested and mown (21 and 27 days, respectively). In the third and final campaign, all three blocks were harvested. Consequently, for the final data collection campaign, there were three different growth periods (7, 14 and 34 days).

Each plot was randomly allocated a different nitrogen (N) fertilizing regime (0, 25, 50, 75 or 100 kg N/ha). The fertilizer was manually applied (i.e., top-dressing) on each plot at the start of each regrowth cycle, having urea as N source. Prior to spring (second half of August) and prior to installing the experiment, phosphorus (P), potassium (K) and sulfur (S) were broadcast throughout the trial area according to soil analysis to ensure that the lack of macronutrients would not impede plant growth.

Data collection campaigns consisted of three subsequent stages: (1) multispectral imagery acquisition, (2) proximal hyperspectral measurements, and (3) biomass determination (Figure 5.1 - Data Collection Workflow). Spectral data was collected by a field spectrometer (ASD Handheld 2 or ASD FieldSpec 4, CO, Boulder, USA, in 2017 and 2018, respectively) and a UAV mounted multispectral camera (Parrot Sequoia - Ile-de-France, France, Figure 5.1 - Instruments). Four collection campaigns were performed under clear-sky

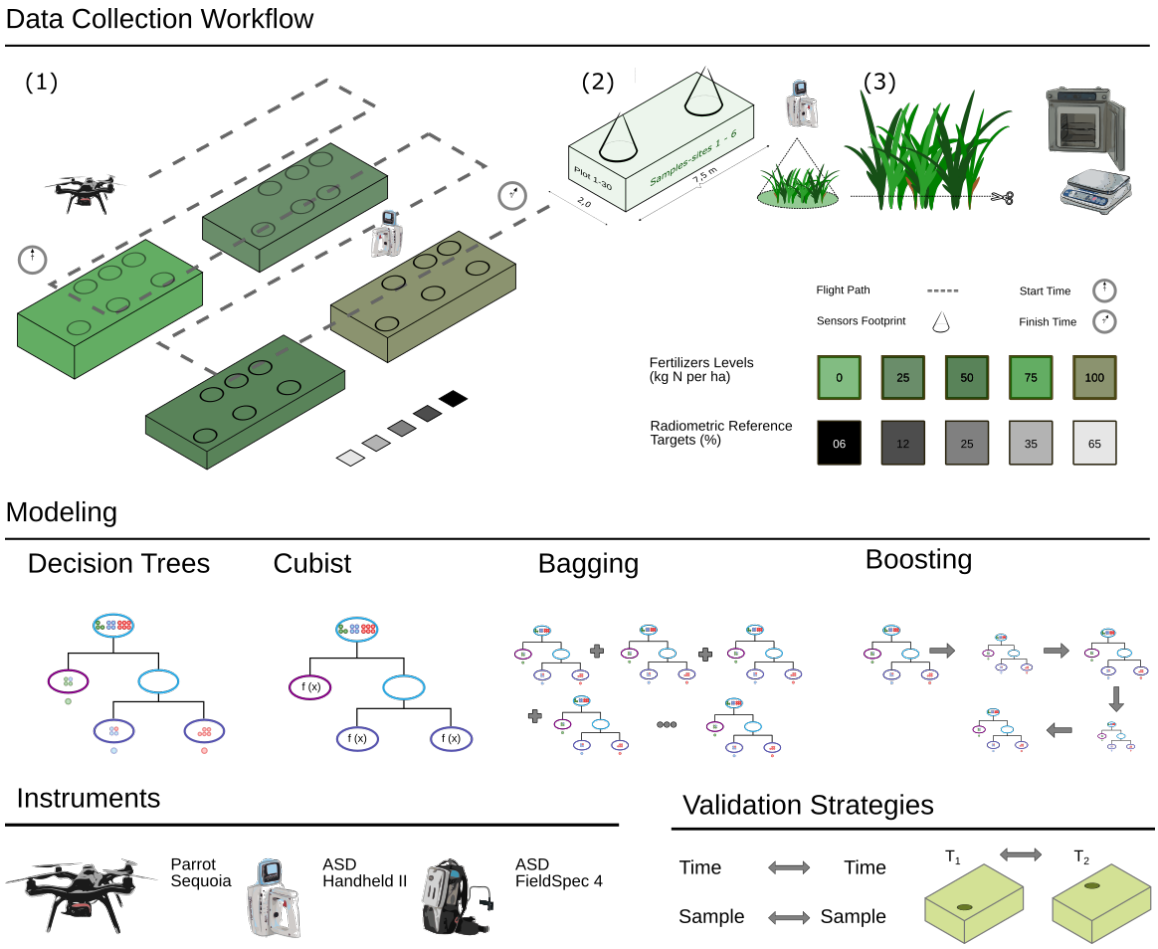



**Figure 5.1:** Workflow for data collection (1 – Mission Planning, 2 – Data Collection, and 3 – Ground-Truthing). Modelling (Regression Techniques). Instruments and Validation Strategies (cross-validation and temporal-validation)

conditions and around solar noon: November 28$^{th}$, 2017 and the 11$^{th}$, 17$^{th}$, and 24$^{th}$, November of 2018.

For the proximal measurements, no fore-optics were attached to the field spectrometers. Both instruments displayed an equivalent field of view (FOV - 25°), and a 1 nm bandwidth across the 400–1100 nm spectral range. Total time spent to obtain all spectral measurements (180 data points) ranged from 1.5 to 2 hours per field campaign with minimum warm-up of 30 minutes. The instrument setup follows the manufacturer's recommendation: 30 scans for spectrum averaging, 60 scans for dark current and white reference.

Within each plot, six sample-sites were randomly chosen. Before the start of the data collection campaign, markers were placed next to each sample-site. Each sample-site was measured five times. The final sample spectral value was the average value of these five measurements. The sequence of measured plots was randomized to minimize any systematic effect of solar position across the plots during data collection. In addition,

after finishing measuring the samples of each plot, a spectral measurement of the white reference (Spectralon®) was recorded.

The intention of this procedure was twofold: (a) to monitor the stability of the instrument and (b) detect any possible change in atmospheric conditions. The instrument was recalibrated (against the white reference) after seven minutes of continuous usage or whenever the white-reference measurement deviated from 100% reflectance, whichever occurred first. Within each plot, six randomly allocated sample-sites were selected. Proximal spectral measurements were taken from approximately one-meter height, thus, yielding a circular footprint equal to 0.15 m (or 0.44 m diameter). Each sample-site was measured five times. Final sample spectral value was the average value of these five measurements.

Pasture biomass was mechanically defoliated above a residual height of 50 mm from the 0.15 $m^2$ spectrometer footprint. Harvested material was dried for a minimum of 48 hours at 60 °C in a forced-air oven immediately following each harvest. Samples *dry-matter (DM)* were weighed using a digital scale (MassCal, 30 kg ±0.5 g). In total, 480 biomass-samples were collected.

*Mission Planning*

UAV flights path ensured a minimal of 75% overlap between images, using a survey-grid flight path, flight-height of 35 m and speed of 3 m/s, with a north-south flight-track orientation, and a camera time-lapse interval set to one second. Flights were performed within ±1 hour of solar-noon, following the procedure detailed in Fallet and Domenzain (2018).

Prior to the flight, the multispectral camera settings were optimized through the camera's web interface, using the spectral calibration target provided by the manufacturer. The total area surveyed was 0.1 ha (25 x40 m), and flight characteristics ensured a ground-sampling distance below 5 cm. Flight duration was less than five minutes. No warm-up period was used to avoid thermal instability and dark current effects.

Five aluminum-made reference panels with dimensions 0.4 x 0.4 m, were laid out next to the experimental area. These were painted using different gray level coatings from off-the-shelf Mankiewicz Nextel® (Hamburg, Germany) Suede Coatings 3101 (Anthracite, Stone Gray, Light Gray, Pearl and Cream) with average reflectance values from 6% to 65% (across the 400 –900 nm range). The reflectance values of these panels were measured in laboratory using a Malvern PanAnalytical FieldSpec 4 and a contact probe taking measurements across the panels.

Ground control points were acquired for each plot, using a RTK GNSS (Reach RS - Emlid, Saint-Petersburg, Russia). Yet, provided the large number of images in a flat small area,

the photogrammetric (SFM) process was sufficient for accurate corregistration of bands and spatial accuracy without the use of ground control points (Pricope et al., 2019).

*Imagery Radiometric Correction.*

Imagery was processed in Agisoft Metashape Pro (Agisoft LLC, St. Petersburg, Russia) following the guidelines provided in Fraser and Congalton (2018) and Agisoft (2020). The raw imagery was radiometrically corrected using both the irradiance sensor and the known reflectance targets (i.e., the panel provided by the manufacturer plus the five reference panels - Method "B" of Poncet et al. (2019), yet employing additional reference panels). The orthomosaic was generated using the *Mosaic Mode* as `blending mode` and pixel size to 5 cm. This method is refered to as the *Pre-Calibration* method.

A second and subsequent correction was performed using the known reflectance values for each of the reference panels, following a modifiction of the method described in Poncet et al. (2019) (Method "D"). This *Post-Calibration* method consisted in applying a linear function to the orthomosaic. This linear function was derived from the average reflectance of the known reflectance targets as displayed in the orthomosaic and their (true) measured value in laboratory.

### 5.2.2 Data Analysis

Data analysis was performed in RStudio/R (versions 1.2.5 and 4.0.2, respectively). For reproducibility purposes, data analysis operations are introduced by the corresponding `package::function` format (i.e., typewriter typeface and accompanied by the double colon operator, i.e. the scope resolution operator).

### 5.2.3 Reference Observations

To ensure comparability between models, a *post-hoc* analysis, the Dunn's Multiple Comparison Test (`FSA::dunnTest`) was performed over the biomass observations, having *Year* as the comparison factor. Additionally, a Principal Component Analysis (PCA) was performed over the spectral benchmark responses (i.e. those collected using the handheld spectrometer) and *Year* was used as the attribute for comparison (`factoextra::fviz_pca_ind`). Both analyses serve the purpose of testing whether biomass distributions and spectral observations from different years (i.e. 2017 and 2018) were comparable.

**Spectral Accuracy Assessment** - Hyperspectral data was convolved (`hsdar::spectralResampling`) to the same specifications as the commercial multispectral sensor . This camera has four dedicated imaging sensors, corresponding to the green (B1, 530–570 nm), red (B2, 640–680 nm), red-edge (B3, 730–740 nm) and near-infrared (B4, 770–810 nm), an irradiance meter compatible with the imaging sensors, and provides a spectral

calibration target of known-reflectance target for camera optimization and radiometric correction. The spectral resolution and the spectral responsivity were extracted from the manufacturer's technical sheet using WebPlotDigitizer (Rohantgi, 2020).

For the multispectral imagery, polygon layers were generated to match the location of all handheld point measurements through a parametric script. This script employed the final output of the randomly generated location for handheld measurement. The accuracy of the polygon layers was checked against the markers (visible in the imagery) and its known location within the plot. These polygon layers were then used to extract (`raster::extract`) the spectral data related to the sensor footprint ($n = 720$).

### 5.2.4 Biomass Modelling

Five different regression algorithms were employed: Classification and Regression Trees (CART) (Breiman et al., 2017), Cubist (Quinlan, 1992), Bagged Trees (Breiman, 1994), Boosted Trees (Schapire, 1990), and Random Forest (Breiman, 1984). These models share the same decision-rule architecture, employing strategies that are able to progressively map more complex fits, while employing techniques for balancing both model bias and variance. However, these techniques render an increasing level of complexity to each model, reducing the interpretability.

Both CART and Cubist are more interpretable than the remaining models. CART develops a regression-tree in which its end-leaf is computed as the mean of the observations selected through the rules associated with that branch. With a slight modification, Cubist computes linear-functions rather than averages; thus, each end-leaf is a piecewise linear model (Figure 5.1 - Modelling. Decision Trees and Cubist).

Both bagging and boosting are ensemble techniques (Opitz and Maclin, 1999). The concept of boosting refers to the weighted sum of a one-level regression tree (i.e., one predictor regression tree, also known as a "weak learner") to create an sequential and additive model, further improving (i.e., boosting) model performance (i.e. "strong learner"). The concept of bagging (bootstrap aggregating, also refers to a meta-learner concept), explore the use of many models generated through resampling with replacement (i.e, bootstrap), the prediction is weighted average (i.e., aggregate) of all models (Figure 5.1 - Modelling. Bagging). The bagging prediction is, thus, an average results of large number models.

The model performance was evaluated using a nonparametric pairwise multiple comparisons *post-hoc* test: Dunn's test (Dinno, 2015). This analysis provides an insight into whether different algorithms have performed significant different from others, while grouping algorithm of similar performance.

Similarly to the regression trees, the Cubist algorithm may also employ a technique similar to boosting named "committees", by which several model trees are created in sequence. Also, a smoothing technique named as "instances" may be employed to further improve

model performance (Quinlan, 1993). The trade-off between these additional techniques and interpretability can be taken into account when opting for a final model.

The predictors employed for modelling were: *Normalized Difference Vegetation Index (NDVI)* , *Green Normalized Difference Vegetation Index (GNDVI)*, *Transformed Vegetation Index (TVI)*, *Normalized Difference Red Edge Index (NDRE)*, *Leaf Chlorophyll Index (LCI)* and *Modified Chlorophyll Absorption Ratio Index (MCARI)*, generated as per the equations and references available in `RStoolbox::spectralIndices`.

**Table 5.1:** Vegetation indices employed for biomass modeling

| Index | Formula | Author (Year) |
|---|---|---|
| NDVI | $(B_4 - B_2)/(B_4 + B_2)$ | Rouse et al. (1973) |
| GNDVI | $(B_4 - B_1)/(B_4 + B_1)$ | Gitelson et al. (1996) |
| NDRE | $(B_4 - B_3)/(B_4 + B_3)$ | Haas et al. (1975) |
| TVI | $\sqrt{(B_4 - B_2)/(B_4 + B_2) + 0.5)}$ | Gitelson and Merzlyak (1994) |
| LCI | $(B_4 - B_3)/(B_4 + B_2)$ | Zebarth et al. (2002) |
| MCARI | $(B_3 - B_2) - (B_3 - B_1) * (B_3/B_2)$ | Daughtry (2000) |

All the indices employed (Table 5.1) are compatible with the bands available in the multispectral camera and commonly employed for precision agriculture purposes (Franzini et al., 2019; Lu et al., 2020).

### 5.2.5 Validation Strategies

Validation strategies allow for insights in model performance under different circumstances, evaluating the applicability of boundary conditions of the models training stage. Two different validation strategies were employed to measure performance across different scenarios: (i) repeated $k$-fold cross-validation (Stone, 1974); and (ii) temporal-validation (Figure 5.1 - Validation Strategies).

The $k$-fold cross-validation provides an insight of baseline error whenever the boundary conditions for both model training and testing are the same. Temporal-validation provides insight regarding the performance of models trained in one year and validated in a different one. Thus, in the context of this study, two training and validation sets were available (2017 and 2018) and were used alternately.

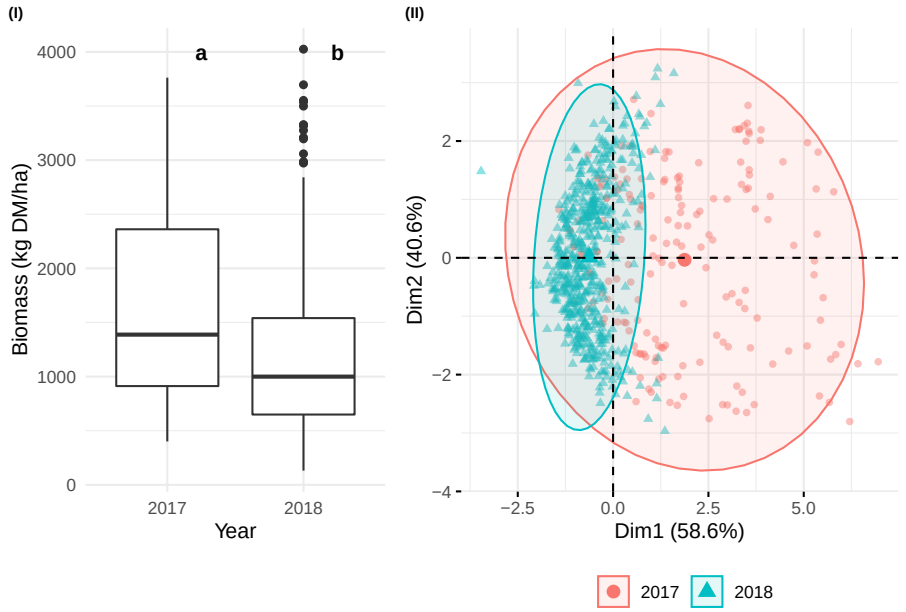



**Figure 5.2:** Reference Observation - Boxplot biomass values per year (I) and PCA of spectral reference ASD (II); the ellipses represents 90% of the corresponding observations.

## 5.3 Results

### 5.3.1 Reference Observations

**Biomass observations** - For the complete data set, observations ranged from 132 to 4025 kg DM/ha, with an average value of 1356 and median 1125 kg DM/ha (Figure 5.2 - I).

Both years presented similar ranges, however, Dunn's test indicated that distribution between both years were not comparable. The biomass values observed in 2017 presented a broader distribution of quartiles (mean = 1646, maximum = 3762 , minimum = 401 kg DM/ha), in comparison the 2018 observations (mean = 1182, maximum = 4025 , minimum = 132 kg DM/ha), as seen in Figure 5.2 - I.

**Spectral observations** - The results of the PCA were able to summarize 99.2% of total variance of the original data set in the two first Principal Components (Biplot. Figure 5.2 - II). To a large extent, the PCA indicates that the spectral observations were dissimilar: observations for both years are clearly projected in different areas of the biplot, as shown in (Figure 5.2 - II ). As a reduced feature space, the PCA biplot indicated the 2017 data displayed a wider spectral variability than 2018, encircling observations of the first year within both Principal Components.

### 5.3.2 Imagery Radiometric Correction

**Pre-Calibration** - Following the radiometric calibration performed exclusively through Agisoft Photoscan (i.e. Method "B"), the analysis of the radiometric response for reference targets showed a non-linear response in all four bands (Figure 5.3 - Pre-Calibration). Particularly, a complete saturation of the Green and Red bands (B1 and B2, respectively) was evident past the range of 20-30% reflectance. In addition, the multispectral camera did not consistently under- or overestimate the reflectance values, displaying no systematic error across dates (Figure 5.3 - Pre-Calibration). Such is made explicit through the through the comparison of 1:1 line between spectrometer and multispectral camera. In different dates and for different bands, the camera has displayed slopes which were either smaller or larger than one.

**Post-Calibration** - Following the methodology presented in Poncet et al. (2019), the reflectance values were further corrected through a linear regression against the reference targets. Due to the saturation displayed for both the Green and Red Bands, only the the darker targets (i.e. Dark-Anthracite, Stone-Gray and Light-Gray) were employed as data points in the empirical line calibration(Figure 5.3). For the remaining bands, all targets were used in the linear regression (Figure 5.3).

Without taking into account the errors due to saturation of the Green and Red Bands (B1 and B2, respectively), the average RMSE for all bands across all dates was equal to 5.89% and 1.28% for pre and post-calibration, respectively (Figure 5.3).

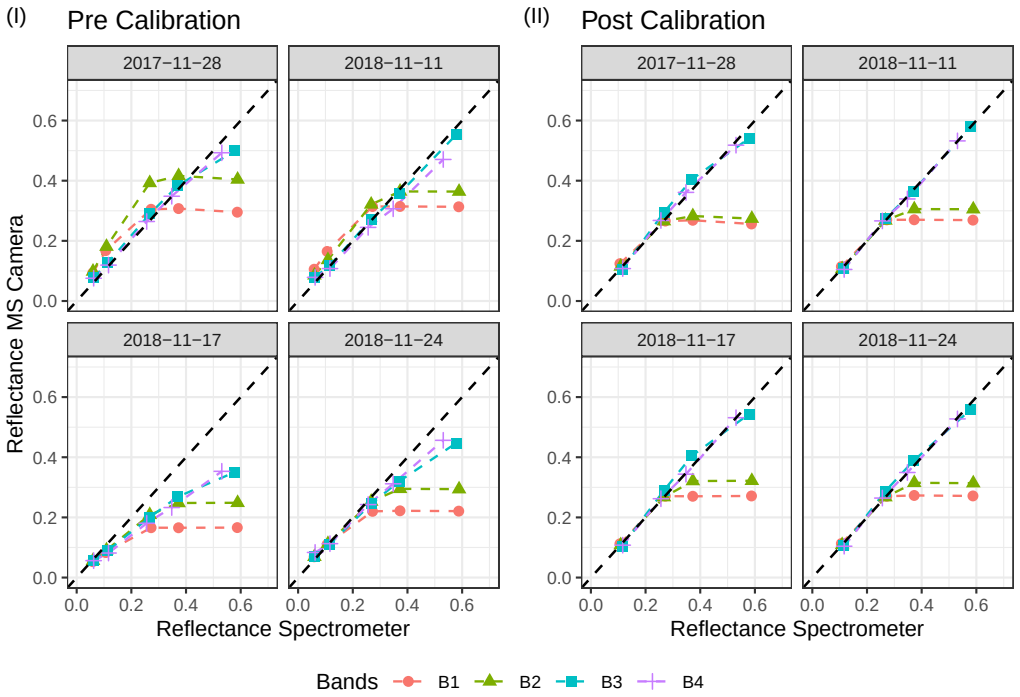



**Figure 5.3:** Scatterplot of Multispectral Camera and Handheld Sensor Measurements of Reference Panels Spectral. Pre Calibration refers to the radiometric calibration as per the manufacturer's guidelines. Post Calibration refers to the Method "D" described in Poncet et al. (2019)

**Radiometric Assessment** - An assessment between the handheld spectrometer (i.e. benchmark) and the corrected orthomosaic ($n = 720$) is presented in the Table 5.2. For all dates and bands the determination coefficient was below 0.90. In absolute values, the range of *root mean square error* (RMSE) was equal to 1% – 19%. The *normalized root mean square error* (NRMSE), express the RMSE value as a fraction of the range of reflectance values for each band. The range of reflectance measurements per date (maximum, mean and minimum) are presented in Figure 5.4.

Particularly for the date of November 28$^{\text{th}}$ 2017, the difference between handheld measurements and spectral imagery is found to be the largest. The date with best agreement between sensors is found in 17$^{\text{th}}$ November 2017. Across the bands, the red-edge (B3) is the band with lowest correlation between both sensors, with maximum and minimum determination coefficients equal to 0.53 and 0.16, respectively. The highest level of error is found for the near-infrared (B4), reaching 19% in the 27$^{\text{th}}$ November, 2017.

Overall, all bands displayed a poor performance, presenting a median NRMSE equal to 19% Figure 5.4.

**Table 5.2:** Correlation and Error Metrics between handheld sensor and post-calibrated multispectral camera

| | $R^2$ | | | | RMSE | | | | NRMSE | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 28 Nov | 11 Nov | 17 Nov | 24 Nov | 28 Nov | 11 Nov | 17 Nov | 24 Nov | 28 Nov | 11 Nov | 17 Nov | 24 Nov |
| 550 | 0.19 | 0.30 | 0.75 | 0.59 | 0.03 | 0.03 | 0.01 | 0.01 | 20.4 | 20.7 | 13.2 | 21.8 |
| 660 | 0.64 | 0.48 | 0.86 | 0.63 | 0.03 | 0.03 | 0.01 | 0.01 | 23.0 | 21.7 | 11.2 | 15.3 |
| 735 | 0.16 | 0.23 | 0.51 | 0.53 | 0.14 | 0.03 | 0.02 | 0.03 | 105.7 | 18.8 | 13.5 | 19.1 |
| 790 | 0.28 | 0.56 | 0.80 | 0.78 | 0.19 | 0.04 | 0.04 | 0.06 | 71.9 | 18.5 | 18.2 | 16.5 |

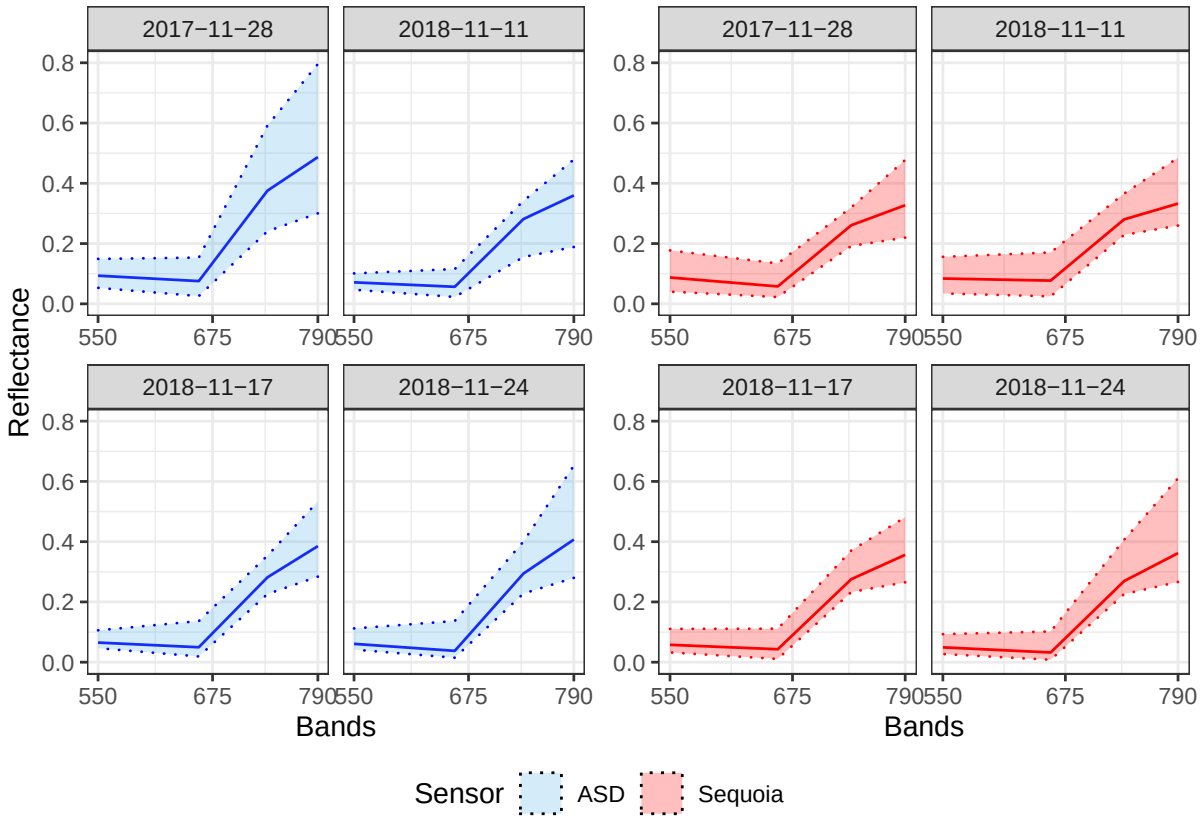



**Figure 5.4:** Average, Maximum and Minimum Spectral Response per Date and per Instrument. ASD Fieldspec spectral response (left-side, blue). Parrot Sequoia spectral response (right-side, red). The coefficient of determination ($R^2$), RMSE and NRMSE, between both instrument is presented in Table 5.2.

### 5.3.3 Biomass Modelling

The results of *k*-fold cross validation are presented in Figure 5.5. Dunn's Test were used to identify which specific means are significant from the others (i.e. different *groups*).

**Multispectral camera** - for the multispectral camera dataset (Figure 5.5 - I), within the five different models, both Cubist and Random-Forest have the lowest and equivalent

**Table 5.3:** Model sizes and Prediction Time (based on the raster file from November 28$^{th}$, 2018)

| | CART | Cubist | Bagged Trees | Boosted Trees | Random Forest |
|---|---|---|---|---|---|
| Time (s) | 5.0 | 55.1 | 5.8 | 8.0 | 18.1 |
| Size (Kb) | 1033.6 | 98.5 | 4244.2 | 1746.7 | 5706.9 |

error distributions as per the results of Dunn's test (Figure 5.5 - Group b) with an average RMSE of 397 kg DM/ha. The three remaining regression algorithms, bagged-trees, boosted trees and CART (Group a), presented a similar average RMSE of 441.5, 446.4 and 463.4 kg DM/ha, respectively.

Both CART, bagged trees and random forest have presented a bi-modal distribution. Conversely, both Cubist and Boosted Trees displayed an unimodal distribution of errors. The difference between Groups *a* and *b* is equivalent to 53.3 kg DM/ha.

**Handheld sensor** - when the handheld spectral data was employed (Figure 5.5 - II), the best performing regression algorithm was Cubist (464.2 kg DM/ha), followed by Random-Forest (519.5 kg DM/ha), Bagged Trees (540.9 kg DM/ha), Boosted Trees (572.8 kg DM/ha) and CART (613 kg DM/ha).

As per the Dunn's Test results, only Cubist (Group b) do not share a similar error-distribution with any other algorithm. Both Random Forest (Group c) and Bagged Trees are equivalent. Similarly, both Bagged Trees and Boosted Tree (Group d) are equivalent as well as Boosted Trees and CART (Group a).

Overall, for the best performing regression algorithm was Cubist and the differences in average RMSE for both sensors was equal to 66.9 kg DM/ha. Also, the multispectral camera presented lower levels of error than the handheld spectral sensor.

**Model Size and Prediction Time** - When deployed for prediction, using a off-the-shelf laptop (Dell Precision 3530 - Windows 10) the fastest regression algorithm was CART and the slowest was Cubist (Table 5.3), taking nearly 11 times more time to fulfill the same task. When compared against the second slowest algorithm (i.e., Random-Forest), CART was 4 times faster. In terms of model size, the smallest model was Cubist (98.5 kB) and the largest Random-Forest (5706.9 kB).

**Model Tuning and Selection** - All models hyperparameters were tuned based on the set of parameters for the minimal RMSE. In the best performing model (i.e., Cubist) both the number of 'committees' and 'instances' were optimized. For the multispectral camera, minimal RMSE was found using nine instances and ten committees (Figure 5.6 -

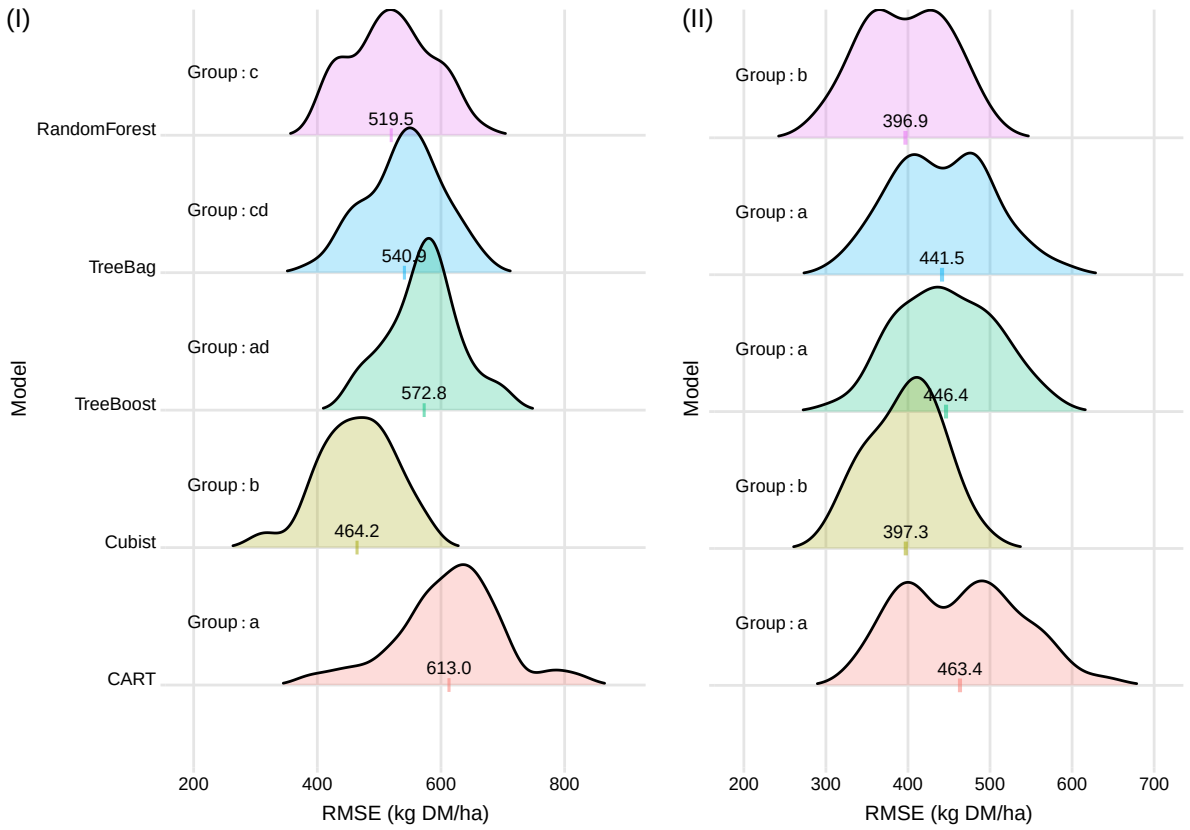



**Figure 5.5:** Root Mean Square Error (RMSE) distribution for $k$-fold cross validation: (I) error-distribution for using a handheld spectrometer; (II) error-distribution using the multispectral camera data. A Dunn's test ("Group") was applied for each corresponding sensing instrument to determine whether error distributions were equivalent. Models size and processing times are presented in Table 2

I). However, without the use of a committee or instances, the mean RMSE was equal to 428.6 kg DM/ha, a decrease of 31.3 kg DM/ha.

The same optimal tuning characteristics were found for the handheld measurements (i.e, nine instances and ten committees). However, without the use of committees or instances the decrease in performance is smaller (12.3 kg DM/ha, Figure 5.6 - II).

The use of committees had a larger effect on model performance than instances, indicating that sequential models (i.e., boosting techniques) focusing in poorly predicted observations were beneficial for model performance.

**Model Interpretability** - A Cubist regression-tree was generated, as an example of model interpretability using the its M5 version (`RWeka::M5P`), for both the handheld and multispectral camera data sets (Figure 5.7). For the purpose of brevity, only the Figure concerning the multispectral camera is presented, while the handheld is provided as an

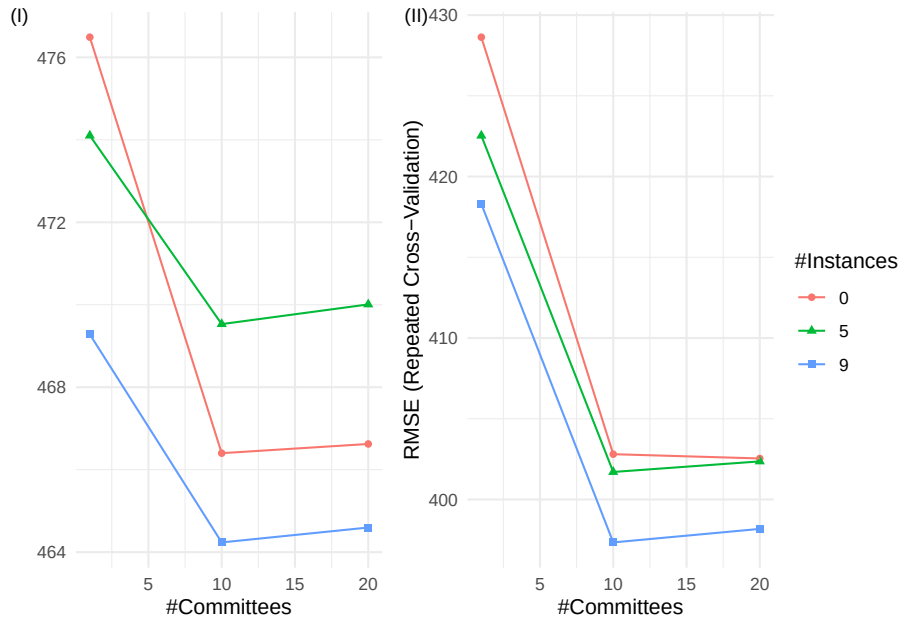


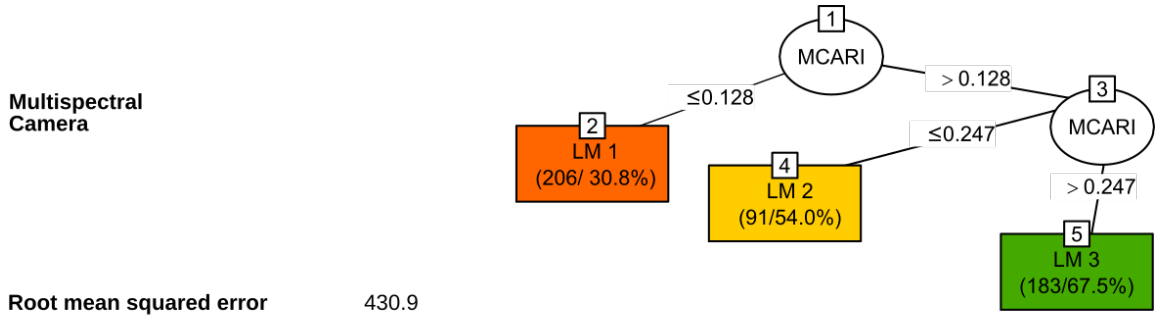

**Figure 5.6:** Hyperparameter tuning for the Cubist regression algorithm. Sub-figures I and II corresponding to the model tuning using the handheld and the multispectral camera, respectively.

annex. The performance of the algorithm was equivalent to the values found within the tuning of hyperparameters (i.e., lowest values of committees and instances). For both sensors the number of splits were small, the first rule-based split used on the MCARI and the worst performing group was that associated with a low response in MCARI. Particularly for the handheld sensor, additional splits were based on NDVI and NDRE. Noticeably, for all *linear models* (LMs), all coefficient values are high, possibly indicating overfit within each model.

**Temporal Validation** - In the temporal validation strategy, Cubist was the only model assessed as all other algorithms had been found sub-optimal through the repeated $k$-fold cross-validation, thus, rendering an extensive analysis futile.

Regardless of the sensor, predictions were always highly biased. The prediction results of 2017 were underestimated, while the results for 2018 were over-estimated (Figure 5.8). When the validation set was the 2017 year, the RMSE was equal to 1087 and 797 kg DM/ha, for the handheld and multispectral camera, respectively. Conversely, for the year of 2018 the RMSE was equal to 1258 and 1432 kg DM/ha, respectively.

Regarding the stability of the error-metrics, the handheld sensors were shown to be more consistent, with a difference between years equals to 171 kg DM/ha, whereas the multispectral camera has a wider variation between both years (635 kg DM/ha).




**LM 1.** Biomass = -1089.5948 * GNDVI + 2386.7631 * MCARI + 6971.5858 * NDRE + 649.5145
**LM 2.** Biomass = -1066.1053 * GNDVI + 240.9809 * MCARI + 3030.2105 * NDRE + 1604.102
**LM 3.** Biomass = -8242.4036 * GNDVI + 1139.2886 * MCARI + 18240.1298 * NDRE + 4719.1124

**Figure 5.7:** Cubist Regression (multispectral camera). Terminal leaves present both the number of observation and coefficient of determination for its respective linear model (LM)

Interestingly, across sensors, the bias introduced by either under or overestimation inflated RMSE values, generating results that have comparatively higher RMSE and higher coefficient of determination. This is noticeable through the visual inspection of more compact or elongated data point-clouds which, although having higher RMSE, presented a higher coefficient of determination ($R^2$ - Figure 5.8).

To allow a better comparison of model performance, a third statistical error-metric was added: the Concordance Correlation Coefficient (CCC) (Lin, 1989). Through it, we can infer that none of these models has an optimal performance (Figure 5.8). All models presented a CCC lower than 0.5. Under this metric, however, the model associated with the inputs of the multispectral camera displayed a stable metric (CCC $\approx 0.45 - 0.48$), while not displaying an accentuated level of saturation at high biomass-levels.

## 5.4 Discussion

This research aimed to assess the (i) equivalence between spectral data acquired at top-of-the-canopy and low-level flight, (ii) evaluate the impact of increasingly complex tree-based models, and (iii) estimate model performance under a repeated $k$-fold cross-validation and a temporal validation. Our results have shown that a thorough examination of the radiometric quality of the multispectral camera is necessary. Firstly, the camera has clearly indicated a response-saturation both in the Green and Red bands at levels above 20-30% reflectance (Figure 5.3). Consequently, the camera is limited to spectral observations of vegetation, for which most targets are within such range (i.e., reflectance below 20–30%).

Additionally, although our experimental design did not aim to measure it, the analysis (Figure 5.3) indicates a lack of dynamic range of the multispectral camera in comparison

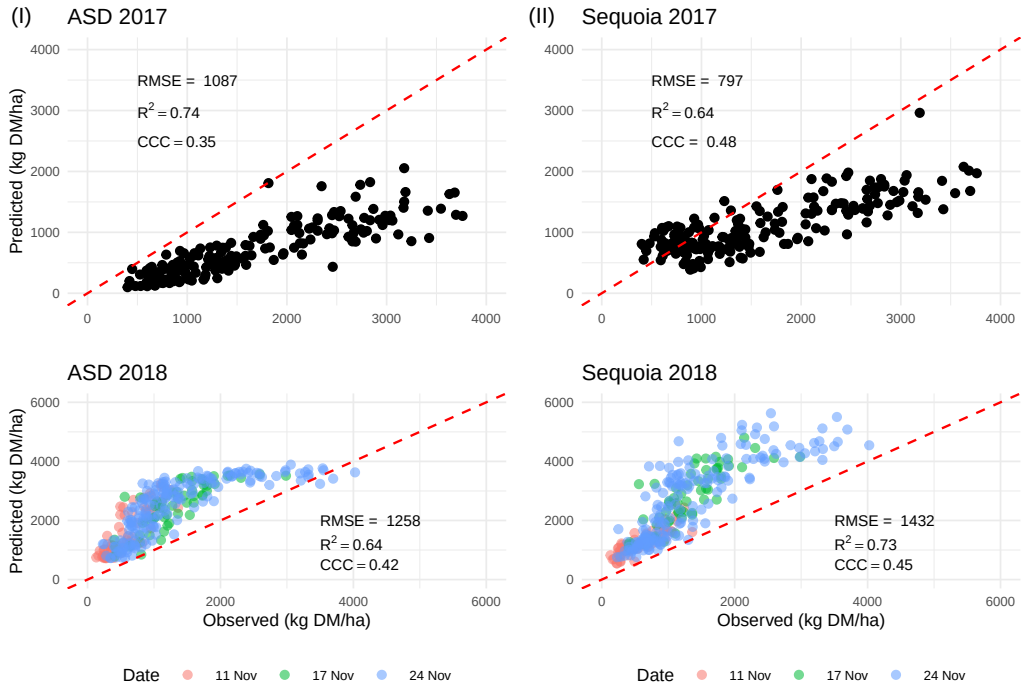



**Figure 5.8:** Predictions for temporal validation strategy. Scatterplots on the left and right side, present the validation results for the handheld and camera data, respectively. Accordingly, the top and bottow row refer to the years used as validation sets, either 2017 or 2018, respectively

to the handheld sensor (Figure 5.4). This can be seen in wider amplitude of values present in the near-infrared band (B4) of the handheld sensor in comparison to the camera. Furthermore, the red-edge band (B3) has shown little correlation with the measurements made by the handheld sensor (Table 5.2. Finally, as the median NRMSE (19%), shows that the multispectral camera has a poor performance in its operational range.

Although the method defined in Poncet et al. (2019) was able to correct the panel spectral responses, reducing the average RMSE to a fifth of its initial value, the spectral observations across the orthomosaic presented a poor correlation between the handheld measurements and the multispectral imagery. The source of this error cannot be elucidated based on the information available, but is likely to be related to several shortcomings of the multispectral camera: (a) absence of a open-source radiometric correction pipeline, where researchers could check, adjust and improve the workflow; (b) the irradiance sensor has a single and fixed field-of-view, which may not be sufficient to provide suitable tilt-correction as discussed in Suomalainen et al. (2018). As is, the radiometric correction workflow based solely on the irradiance meter and reflectance-target provided by the manufacturer is most likely not sufficient for accurate quantitative purposes.

Regarding model interpretability and performance (Carvalho et al., 2019), the Cubist algorithm has shown to be a satisfactory option, outperforming other methods evaluated.

Although its speed has not been satisfactory, this can be attributed to the implementation in the R-language. The original versions of the algorithm were coded in the C++ language, resulting in a fast implementation which should outperform other models. A useful proxy for its speed is model-size, in which Cubist was the smallest. Additionally, as seen in the Figure 5.7, these models are far from complex and resemble a CART algorithm, consequently, its deployment should be optimal in an operational scenario. Furthermore, its minimal size may also prove itself advantageous as an embedded technology in low-cost, low-processing power equipment such as handheld devices or embedded in agricultural machinery.

An additional desirable characteristic of these models is its interpretability. From Figure 5.7, one can identify terminal-leaves with poor performance (e.g. observations with low MCARI responses) and further investigate the reasons for such. In comparison with black-box models, Cubist has great potential as a simple technique and able to provide insights about the inner-workings of the model.

The idea of boosting through committees, may prove itself worthwhile, if the subsequent-level boosted tree can be examined, analogous to the analysis of subsequent components or latent structures in *Principal Component Analysis* and *Partial Least Squares Regression*, respectively. The *linear models* within Cubist may benefit from feature selection and regularization, as high values of model coefficients indicated overfitting or multicollinearity, which would further improve the model's interpretability.

The ability to accurately predict a different year was also evaluated, showing poor results. This is most likely due to the non-overlapping characteristics of both biomass and spectral observations, as presented in Figure 5.2. Such results, indicate the necessity of a wide training set of observations both spectrally and in the biomass domains. Also, due to the large bias, both in under and over-estimation of errors, the analysis and deployment of the models presented should be carefully evaluated and most likely avoided. This, however, shows the shortcomings of reporting achievable accuracies in short-term studies.

Regarding the performance of instruments, the post-calibrated multispectral camera has presented a better performance than the handheld sensor. Several factors may have contributed to such: (i) the non-optimal directional response function (DRF) associated with both the ASD Handheld and FieldSpec, which is non-homogeneous across the sensors footprint (Mac Arthur et al., 2007), (ii) the averaging of the bidirectional reflectance distribution function introduced by Agisoft `blending mode`; (iii) the smaller data collection interval from the multispectral camera, consequently controlling for differences in illumination geometry.

It is important to stress that these results were possible due to the introduction of additional known reference panels and further calibration of reflectance values. In this sense, given the low level of correlation found in the *pre-calibration* stage, it seems reasonable to advocate

for the use of additional reference targets whenever radiometric accuracy is a relevant factor.

The absence of a open workflow for radiometric correction, from digital number to radiance and reflectance measurements is a significant handicap. As a consequence, end-user are left without any form of data quality-assurance or trust-worthiness of the end-products. Consequently, laborious procedures, such as those described in Poncet et al. (2019), are required to ameliorate radiometric accuracy in spite that these measurements and procedures are embedded in the manufacturer's pipeline.

The absence of consistent correlation levels with handheld measurements, if introduced by poor camera performance, will significantly decrease the potential of promising techniques for biomass assessment, such as time-series analysis. Since our data-collection campaigns, a new irradiance-meter has been introduced by the manufacturer, in a similar design as the one introduced in Suomalainen et al. (2018), which may have improved the radiometric quality of mosaics processed through the commercial pipeline. Provided that the irradiance meter is continuously collecting data during the mission, it may prove itself worthwhile to provide "irradiance" map, so possible measurements error may be corrected or the illumination conditions may be examined *a-posteriori*, similarly to our white-reference scans in our handheld data-collection protocol.

The saturation of the Green and Red Bands, may have been intentional, allowing an exclusive and wider-dynamic range for reflectance levels which are commonly found for vegetation. However, the absence of control of the dynamic-range prevents the full assertion of this claim, while also limiting the use as a scientific-instrument for target with a wide range of values in the Red and Green Bands. The shortcoming of this sensor have also been discussed in Franzini et al. (2019).

The performance of spectral models when assessed through the temporal-validation strategy was likely handicapped by the distinct training sets, as shown in PCA biplot (Figure 5.2 - (II)). Alternatively, the $k$-fold cross-validation illustrates achievable accuracies and similar results as previous studies (Karunaratne et al., 2019; Togeiro de Alckmin et al., 2020a)

Finally, data collection employing a multispectral camera and UAV were considerably faster (i.e. five minutes) than the use of a handheld sensor (one and half to two hours), providing a clear evidence of the usefulness of unmanned platforms for data collection.

## 5.5 Conclusion

This study successfully compared matching spectral measurements of 480 observations, using both a multispectral camera and handheld sensors. Correlation levels between UAV multispectral camera and handheld sensors was not satisfactory, presenting low-levels of correlation ($R^2$). Yet, after radiometric *post-correction*, model performances were similar,

with an average RMSE ranging from approximately 400 to 460 kg{} DM/ha for camera and handheld sensors, respectively. Given that the time necessary for data collection using UAV is a fraction than that reliant in manual data-collection, this result indicates a large potential for the use of multispectral cameras at low-level flight for perennial ryegrass biomass estimation.

The workflow for radiometric correction of the multispectral imagery was significantly improved by the method described in Poncet et al. (2019), reducing average RMSE from 5.89% to 1.28%, clearly indicating the need for an improved automated pipeline, making use of additional reference targets. The current proprietary closed pipeline prevents further improvement and, as is, the radiometric quality of orthomosaics solely employing the commercial method is of limited value.

This study has also shown that a simple type of regression-trees algorithm, namely Cubist, has great potential both to minimize the model-prediction error as well as increase model interpretability when compared to other tree-based algorithms. If implemented in a faster processing language (such as C++), it may also prove itself ideal for fast deployment.

Finally, this study has shown model performance inconsistencies when using a temporal-validation strategy, indicating the need for caution evaluation of results based in short-term studies.

Our results and analysis suggest the need for improvements in the radiometric correction workflow for multispectral cameras, both through a higher-level of user ability to set up camera parameters as well as control of the current proprietary pipeline. As is, the lengthy period required for an improved radiometric-calibration of the multispectral imagery offsets the benefits of a short data collection period, as also remarked in (De Rosa et al., 2021).

Nevertheless, the necessary improvements have already been examined (Suomalainen et al., 2018) and are surpassable. If implemented, as our results show, there is strong reason to believe that UAV multispectral cameras can replicate the accuracy achieved by field spectrometers in previous studies, while automating data-collection.

# Chapter 6

# Synthesis

The overarching aim of this thesis is to expand and enable the use of spectral data for perennial ryegrass (*Lolium perenne* L.) management establishing useful links between the techniques of field spectroscopy and optimal imaging systems deployed at low-level flight. Based on this aim, four specific research objectives (Table 1.1) were defined in Chapter 1.

In the first section (*Main Results*) of this chapter, each research objective is addressed based on the results and findings presented in Chapters 2 to 5. These findings are listed and summarized in Table 6.1. In the following section, *Limitations* (Section 6.2) of the thesis are presented and discussed. Section 6.3 provides a *Reflection and Outlook* of future research objectives and farm-applied use of remote sensing techniques for pasture management.

## 6.1 Main Results

The results presented in this thesis are based on data collection campaigns from December 2016 until November 2018, spanning across two different countries (i.e., Australia and the Netherlands) and four experimental sites. The data collection protocol, comprising procedures for spectral observations and reference observations, was identical across all dates and locations. In total, 1416 spectra–biomass sample pairs were collected, and 266 were sent to the laboratory for feed quality analysis. These samples provided a wide range of biophysical and biochemical characteristics, mostly ranging within 10–30% crude protein and 500–3500 kilograms dry-matter per hectare. Such ranges are representative of common values found in an operational farm-scenario.

### 6.1.1 To evaluate the trade-offs between subsets of spectral bands and achievable accuracies for retrieval of biophysical and biochemical parameters. (Objective 1)

*Addressed in Chapters 2, 3 and 4*

Chapter 2 approaches the task of biomass retrieval, expressed as *dry-matter (DM)* (kg DM/ha), through the use of *VIs*. In such task, it was found that a small subset of indices (i.e., four) within the *Visible to Near Infrared Region (VIS–NIR)* region (325–1075 nm) have a nearly equivalent performance to a large number of features (i.e., 20 or more). This equivalency was defined as a ten percent threshold above the minimal *Root Mean Square Error (RMSE)* found using the initial (filtered) subset of vegetation indices (Figure 2.6).

In that chapter, such a satisfactory subset of VIs displayed an average RMSE of around 400 kg DM/ha when paired to the best-performing regression algorithm (i.e., random forests). Additionally, of the 97 vegetation indices employed in this analysis, nearly

80% were found to be redundant, displaying a high level of correlation (i.e., Pearson-correlation $\geq$ |0.95|).

Chapter 3 further developed this point, showing that only 11 out of 161 bands (bandwidth 10 nm, ranging from 400–2500 nm) were needed for optimal retrieval of *crude protein (CP)*, while reaching an RMSE of around 80–100 kg, when expressed in a weight per area basis (kg CP/ha), or 2.5–3.0 % DM, when expressed as a dry-matter fraction. Furthermore, an analysis of variable importance indicated that, in field conditions, the most important spectral bands for crude protein retrieval are located within the red-edge and NIR-shoulder portion of the spectra, rather than the shortwave infrared region (SWIR).

In a complementary fashion, Chapter 4 provides evidence that hyperspectral (within the 550–790 nm range) data can be accurately retrieved from multispectral data (i.e., four bands) with minimal loss in true spectral data and biomass prediction accuracy. Consequently, additional bands should increase the effects of multicollinearity and redundancy of information when employing hyperspectral data for outdoor vegetation assessment, specifically for the retrieval of parameters related to broadband absorption features within this spectral range. Chapter 5 indicates that for the task of biomass retrieval, when employed in a $k$-fold cross-validation, a small subset of VIs reaches the baseline error rate of approximately 400 kg DM/ha (Figure 5.5).

In combination, these studies reinforce the fact that spectral information is an indirect measurement of the underlying biophysical and biochemical drivers of canopy reflectance, as discussed in Verstraete et al. (1996). Consequently, within the VIS–NIR, the accurate measurements of most of these drivers do not rely on a high number nor narrow bands (i.e., narrower than 10 nm). This statement's exceptions would be the bands located close to the red-edge position and the chlorophyll peaks. The retrieval of empirical relationships of key-management parameters is constrained to the relationship between the biochemical and biophysical (e.g., leaf area index) drivers of canopy reflectance and such key-management parameters (e.g., biomass).

**Relationship with existing knowledge** - In the context of spectral analytics, feature (spectral band) selection is a crucial task, yet frequently overlooked, to ensure that optimal accuracies are obtained. Whenever constraints such as sensor manufacturing complexity and data-processing requirements (Price, 1994) pose a significant restriction to its application, feature selection is of significant importance.

Hyperspectral instruments are the ideal scientific equipment to determine the informative part of the spectral for a single and specific retrieval task. However, when addressing the necessary minimum requirements for dedicated sensing systems, multispectral sensors should be considered as a cheaper and less complex option. This rationale is continuously applied for the development of satellite systems, where due to many constraints, it is preferable to limit the number of bands (Price, 1994).

Historically, prior to the development of partial least squares regression (PLSR) (Wold et al., 2001), the most common analytical method for benchtop NIRS was the use of multiple linear regressions. Differently from PLSR, multiple linear regression is highly affected by multicollinearity or redundancy of the explanatory features, limiting the use of the complete NIRS spectra as predictor. At such an early stage, most equipments were also constrained to multispectral measurements due to technological limitations. However, despite such limitations, biochemical components would be successfully estimated using less than ten spectral bands (Fahey et al., 1994; Norris et al., 1976).

In this sense, Kjeldahl and Bro (2010) stated that a common poor-practice is the use of the complete sensor spectral output rather than the informative part of the spectra. In the context of intensive pasture production, Kawamura et al. (2010) and Starks and Brown (2010), discussed the importance of feature selection both for model performance and for less complex sensors. Chapter 3, combines these informations, which to a large extent were sparsely discussed, providing a cohesive study about retrieval of crude protein in outdoor conditions.

In contrast to other studies that rely on techniques such as radiative transfer modeling (Berger et al., 2020) or retrieve crude protein at the leaf level (Curran, 1989), our study indicates that crude protein can be best retrieved when using the VIS–NIR portion of spectra rather than the SWIR. From a chemometric perspective, absorption features for protein-related compounds are usually found above 1100 nm (Curran, 1989). Consequently, causal links are expected to be found above such range. However, in outdoor scenarios, masking and overlapping effects of other components and a lower signal-to-noise ratio may hamper the predictive ability of the SWIR (Gao and Goetz, 1994). Consequently, it may be argued that a better-suited strategy is to retrieve biochemical information through correlated features that display a stronger and wider signal, such as the interaction between canopy properties and chlorophyll content found in the red-edge portion of the spectra.

When selecting a single vegetation index for biomass estimation, a large difference in accuracy can be found, as seen in Table 2.2, and judicious selection of the optimal index is advantageous. The results in Chapter 2 have corroborated the findings of Mutanga and Skidmore (2004a), showing that an optimized index is far superior to commonly employed indices, such as the ubiquitous *Normalized Difference Vegetation Index (NDVI)*. Chapter 4 expands this analysis, indicating that such advantage is negligible whenever a small coherent subset of indices is selected, such as those offered by the bands available in current commercial *Unmanned Aerial Vehicle (UAV)* multispectral sensors and satellite systems (e.g., Sentinel-2).

For the retrieval of CP and biomass, and as indicated by Mutanga (2004), the most informative spectral region found was the red-edge region (680–730 nm). Such region, however, presents shortcomings and accuracy limitations due to the overlapping

characteristics of main broadband absorption features, consequently resulting in an ill-posed problem and ambiguous spectral responses (Baret and Buis, 2008). Similarly to Chapter 4, a set of equivalent techniques for spectral interpolation of the red-edge region are presented in Bonham-Carter (1988), in this case applied to tree species. Such a straightforward spectral retrieval technique indicates that a small number of multispectral bands can accurately describe this region of the spectra. Furthermore, such interpolated spectral region can be used for the retrieval of the red-edge position (REP), as indicated in Dawson and Curran (1998).

Provided that our data collection spanned across two-years and four different locations, it seems reasonable that an achievable accuracy of empirical spectra-based models should be close to 400–600 kg DM/ha for biomass and around 80–100 kg CP/ha or 2.5–3.0 % DM for crude protein. Such ranges have also been found in Karunaratne et al. (2020) and Lamb et al. (2002), although for a smaller number of samples and shorter time periods.

As discussed in the contents chapters, further improvements in accuracy should derive from either distinct modeling strategies or independent complementary data. Price (1994) categorically states that *"for most natural materials 15-25 spectral bands appear to be sufficient to describe spectral variability"*. In agreement with this statement, this thesis has demonstrated that, if a single biophysical or biochemical parameter is to be retrieved, a small number of bands is sufficient to reach the achievable accuracy between spectra and the parameter studied.

**Table 6.1:** Overview of Research Chapters and their main findings

| Ch. | Title | Findings |
|---|---|---|
| **2** | Comparing methods to estimate perennial ryegrass biomass: canopy height and spectral vegetation indices. | - Optimized Normalized Ratio Index performs equally well to the Rising Plate Meter;<br>- Only a small subset of indices is necessary to reach nearly maximum accuracy;<br>- Optimal spectral bands are positioned at the red-edge;<br>- Model stacking improves precision. |
| **3** | Retrieval of crude protein in perennial ryegrass using spectral data at the canopy level. | - For field-conditions, the most informative spectral region lies within 725–1100 nm;<br>- Eleven bands of 10 nm bandwidth are needed for optimal crude protein estimation;<br>- Crude protein can be best estimated when measured in a weight per area basis;<br>- Predictions are assessed against unknown locations, resulting in a minor decrease in accuracy. |
| **4** | Retrieval of hyperspectral information from multispectral data for perennial ryegrass biomass estimation. | - Perennial ryegrass hyperspectral data can be retrieved from sparse multispectral data;<br>- The effective retrieval method is based on a straightforward piecewise function;<br>- Synthetic continuum-removed features and optimized vegetation indices can be generated without loss in accuracy;<br>- Synthetic features and a small subset of vegetation indices have equivalent performance for biomass estimation;<br>- Predictions are assessed in unknown locations, resulted in a minor decrease in accuracy. |
| **5** | Perennial ryegrass biomass retrieval through multispectral UAV data. | - Current commercial standards for radiometric calibration pipelines are not satisfactory;<br>- Radiometric accuracy can be improved through lengthy post-processing and use of known reflectance targets;<br>- Multispectral handheld and RPAS models have similar accuracy;<br>- Straightforward implementation of regression-trees algorithms are accurate, fast-processing and interpretable. |

### 6.1.2 To assess the impact of different regression algorithms for an optimal model fit between spectral features and pasture parameters. (Objective 2)

*Addressed in Chapters 2 and 5*

In Chapter 2, a number of regression algorithm performances were evaluated, namely, Ordinary Linear Regression (*LM*), Multivariate Adaptive Regression Splines (*MARS*), Support Vector Machines (*SVM*), and Random Forests (*RF*). For each algorithm, the distribution of different error-metrics were generated using a repeated $k$-fold cross-validation. The results indicated a significant difference (e.g., paired $t$-tests) between parametric and non-parametric methods (Table 2.5). From these results, non-parametric were able to achieve better results, and, amongst these, non-linear non-parametric methods (i.e., support vector machines and random forests) were found to be the best performing algorithms (Figure 2.8).

In the context of field spectroscopy, the results from Chapter 2 highlighted the key-importance of judicious selection of features, in lieu of an overreliance on regression algorithms. For performance comparison purposes, these algorithm-feature pairs were compared based on average RMSE differences ($\Delta$). Specifically, the use of NDVI resulted in an average RMSE performance difference ($\Delta$RF-NDVI) of approximately 219 kg DM/ha in comparison with RF and the satisfactory subset of four vegetation indices. Such underperformance difference was significantly diminished to ($\Delta$RF-LM) 68 kg DM/ha when the same four VIs were employed within a linear regression.

Complementarily, the performance difference between SVM and RF, two non-linear and non-parametric methods, was of approximately ($\Delta$RF-SVM) 16 kg DM/ha. In summary, this showcases the fact that the judicious choice of a small subset of indices can reduce the RMSE value by nearly ($\Delta$LM-NDVI) 151 kg DM/ha, while the choice of algorithm has an impact of around ($\Delta$RF-LM) 68 kg DM/ha.

Random Forest was found to be an effective regression technique; however it has a few shortcomings: it is a greedy-growing algorithm, growing indefinitely and adjusting for local optima. Consequently, this characteristic leads to a large-size, heavy-processing black-box algorithm. At the core, this algorithm is based on the regression-tree structures, which are (paradoxically) fast, light, and interpretable.

Provided that both Chapters 2 and 4 have shown that a small subset of vegetation indices was sufficient for optimal biomass retrieval, it was reasonable to suggest that gradually more complex regression-trees (e.g., increasing complexity through the use of bagging and boosting), would be able to effectively map the fit between a common set of vegetation indices and biomass. Coherently with this interpretation, the results (Figure 5.5) in Chapter 5 showed that a good performance can be retrieved from a less opaque model

(i.e., Cubist, (Quinlan, 1992)) while presenting additional benefits: smaller size, faster deployment, and a higher level of interpretability.

In conclusion, it seems reasonable to state that non-linear non-parametric methods are an optimal tool for field spectroscopy. A satisfactory fit should be achieved if a suitable pool of spectral bands and representative sample distribution is provided. In the context of multispectral imaging techniques, where large-size data sets should be processed in near real-time, the option of regression-trees, mainly through the Cubist family of regression techniques, seems to be an ideal tool due to its light-weight and high retrieval performance.

**Relationship with existing knowledge** - Our results are in-line with the view proposed by Verrelst et al. (2015), indicating the superiority of non-linear and non-parametric methods for empirical retrieval of biophysical and biochemical parameters through spectral data. Furthermore, it was also shown that the use of the full-spectrum (discussed in Objective 1) is redundant and, consequently, not necessary for an optimal retrieval of biophysical and biochemical parameters.

While Mutanga (2004) advocates for the use of imaging spectroscopy coupled hyperspectral data and neural networks for outdoor mapping of grassland parameters, our results provide evidence that, if limited to a single retrieval task, a small number of features paired with non-linear and non-parametric methods provides more interpretable models of equivalent or superior performance. Tuning of hyperparameters plays a significant role in avoiding overfitting and is further explored through the judicious choice of validation strategies (Objective 3). As shown in Chapter 2, improvements in model performance are mostly due to spectral features rather than regression algorithms.

These results are also in line with the greater scientific attention to "interpretable machine learning" as discussed in Roscher et al. (2020), allowing for model diagnostics and a better understanding of the circumstances where models display a lower performance. As discussed in Kim et al. (2007) and demonstrated in Chapter 5, Cubist provides a robust framework for model visualization and interpretability; while providing great accuracy performance in comparison to other algorithms (Fernández-Delgado et al., 2019).

### 6.1.3 To determine model performances through the use of rigorous validation protocols for different temporal and geographical circumstances. (Objective 3)

*Addressed in Chapters 3, 4 and 5*

Validation strategies are crucial to provide an adequate assessment of model performance under a coherent deployment scenario. In this thesis, three main scenarios have been treated: (i) model performance when applied to similar conditions as those in which the model has been trained, (ii) model performance when applied to new unseen locations,

(iii) model performance when applied to different years. While the first can be assessed through $k$-fold cross-validation, both latter scenarios provide a greater challenge to robust validation.

In this sense, Chapter 3 presents the use of validation strategies for which the validation set is made exclusively of a location that has not participated in model training. These locations differ in country (i.e., Australia and Netherlands) and soil types. Importantly, biomass samples amongst locations were comparable, as indicated per Dunn's test. For the VIS–NIR range, the loss in performance between known-locations (i.e. $k$-fold cross-validation) and unseen locations was marginal, decreasing from 79.2 to 85.5 kg CP/ha, respectively.

In Chapter 4, when applying the same validation rationale for biomass in a two-years observation study, RMSE ranged from nearly 500 to 600 kg DM/ha for known and unseen locations, respectively. Comparatively to Chapter 2, the error (RMSE $\approx$ 400 kg DM/ha) when modeled to a one-year single location study, Chapter 4 indicates a performance loss for different locations and longer duration studies, although still acceptable for practical applications.

In Chapter 5, data was collected in November of 2017 and 2018, employing a handheld spectrometer and a multispectral camera at low-level flight. Such data, however, presented a smaller sample size ($n = 480$) compared to previous studies and dissimilar spectral observations as per the results of a Principal Component Analysis (Figure 5.2). When applying a $k$-fold cross-validation, best results (through the modeling of common VIs and Cubist) were approximately equal to 400 kg DM/ha.

However, when the data was used to assess the performance of models in different years, results displayed a maximum RMSE of 1432 kg DM/ha (employing the multispectral camera and validated in 2018 data). Results were poor for both sensors, and the handheld spectrometer also displayed a poor performance when validated against 2018 data (i.e., 1258 kg DM/ha). It is important to note that benchtop NIRS chemometric models are periodically updated, and that design, optimization and validation is an active area of research (Westad and Marini, 2015). Our results indicate that the same is also of particular importance for field spectroscopy, requiring a broad and diverse number of biophysical and biochemical characteristics that can provide an adequate training set.

**Relationship with existing knowledge** - Although Sanches (2009), Thulin (2008), and Mutanga (2004) have indicated the potential and high-accuracy of field spectroscopy, this thesis aims to evaluate such potential using these cautionary validation strategies.

In the Outlook section of their dissertations, both Sanches (2009) and Thulin (2008) recommend using a large dataset with an appropriate (i.e., unseen) validation sets instead of $k$-fold cross-validation. Our research has pursued this objective and has shown that model accuracy validation can be better assessed by manipulating validation sets by

geographical or temporal factors. Our results are in agreement with the remarks made by Meyer et al. (2019), who claimed that model performance is "*naively*" assessed when employing $k$-fold cross-validation.

Most recent studies (Karunaratne et al., 2020) have not yet incorporated the same precautions. Consequently, it can be argued that rather than a global performance, these studies (Karunaratne et al., 2020; Mutanga, 2004; Sanches, 2009; Thulin, 2008) provide a baseline error or achievable accuracy in *local* scenarios. The same issue is found in chemometrics and laboratory NIRS analysis (Fahey et al., 1994). Instead of a *global* model, different ("*local*") calibration equations (i.e., models with different training datasets) are employed when appropriate to the context that the model will be applied. Consequently, this would require a possibly more extensive geographical range of observations and a temporal update of the model.

These "pitfalls" of class imbalance or improper sampling are studied through experimental design and are not often discussed in remote sensing and field spectroscopy. However, it is crucial to consider the dependency of spectral samples as, usually, these are collected either in the same experimental site (i.e., geographical dependency) or sequential data collection campaign (i.e., temporal dependency). Recently, these issues have been carefully examined, and new validation strategies, such as those employed in this thesis, are becoming more prevalent (Meyer and Pebesma, 2020; Valavi et al., 2019).

### 6.1.4 To explore the transferability of spectra-based models from canopy to low-level flight. (Objective 4)

*Addressed in Chapter 5*

In his dissertation, Pullanagari (2011) explored the use of field spectroscopy for pasture management in New Zealand. In its Outlook, he recommends the *"this study methods need to be extended for space and airborne sensing for regional scale mapping"*. This research objective was addressed in Chapter 5, through the use of UAV, at a scale suitable for farm management. As discussed in the Motivation Section (1.1), the ability to bridge field spectroscopy to low-level flight imaging system can considerably improve monitoring and decision-making for pasture management. Chapter 5 builds on the findings of previous chapters, having a solid understanding of achievable accuracies, the trade-off between spectral features and accuracy as well as validation strategies.

Although spectra-based models both from the UAV multispectral camera and handheld spectrometer have yielded similar accuracies, the radiometric response between both sensors was poorly correlated and inconsistent across dates. This shows a substantial limitation of the multispectral camera employed in this thesis, as other studies have shown that differences between handheld and low-level flight spectral responses can be considered negligible (Von Bueren et al., 2015).

**Relationship with existing knowledge** - Many commercial multispectral systems, particularly low-cost multi-array camera systems (Aasen et al., 2018), are limited in their radiometric accuracy. Specially for the sensor employed in this thesis, other studies have also reported poor radiometric accuracy (Franzini et al., 2019; Poncet et al., 2019). Complementarily, several studies (Mamaghani and Salvaggio, 2019; Poncet et al., 2019) have shown that the standard of-factory pipeline is not optimal for accurate radiometric measurements.

Furthermore, other studies have successfully employed manned airborne spectral observations (Pullanagari et al., 2015; Pullanagari et al., 2016) as well as hyperspectral UAV imaging sensors (Capolupo et al., 2015; Geipel and Korsaeth, 2017; Oliveira et al., 2020) for pasture management. Noteworthy is that most of these studies deploy the use of canopy as explanatory-features. Such has two side-effects: (i) the necessary processing requirement to accurately generate canopy-surface and digital-elevation models at a farm-scale is, currently, prohibitive; (ii) the use of canopy-height leaks information about plant regrowth period, which is a key driver for pasture quantity and quality. Consequently, it is questionable whether model performances are due to either plant-height, spectral information, or regrowth period. Such models have reached equivalent levels of accuracy for both crude protein and biomass, to our studies. Consequently, this thesis provides evidence that such accuracies can be achieved using reliable multispectral cameras, while not requiring extensive processing power (e.g., canopy surface model) nor complementary information (e.g., regrowth period).

In general, this thesis's results indicate that the current commercial pipeline is not sufficient for radiometric calibration. Oliveira et al. (2020) indicate the challenges associated with radiometric correction under unstable environmental illumination conditions while employing a high-end hyperspectral sensor. Such is in agreement with several authors, which have consistently presented improved methods: Suomalainen et al. (2018) demonstrates the importance and new methods to account for sensor tilting, Aasen et al. (2018) discusses the need for corrections due to bidirectional reflectance distribution function (BRDF), topographic and shadow effects under changing illumination (Wang et al., 2019).

Overall, given the results of this thesis and available scientific and technological state-of-the-art, it can be stated that UAV low-level flight remote sensing technology is mature for reliable application. However, best methods and practices have not been widely adopted by sensor manufacturers and end-users. Given the fast development of *UAV* cameras, it is likely that these improvements will be implemented.

## 6.2 Limitations

This thesis would be made stronger if a more detailed error budget had been developed. Although a baseline error (i.e., achievable accuracy) is indicated, this error's sources are not examined in detail. As discussed in Chapter 3, the possibility of spurious correlations between biophysical and biochemical when estimating quality parameters through a weight per area basis has not been thoroughly examined.

This thesis has not examined a larger number of feed quality parameters, which could lead pasture managers to better-informed decisions. However, methods are described in sufficient detail so that this analysis could be further expanded.

Although a significant level of attention has been giving to data collection, this thesis could have profited from a higher number and different multispectral cameras so that the Objective 4 could have been tested further. A possible test for the model transferability of handheld to low-level flight would have been the deployment of models trained with the handheld spectrometer and applied to multispectral imagery sets, which we have not yet performed.

Although aiming to expand the use of spectral data towards imaging systems, we have not addressed the impacts of broader bandwidth and larger pixel-sizes due to a scope decision. Consequently, it is still a research-gap to determine optimal spatial scale and the impact of broader bands for distinct retrieval tasks.

## 6.3 Reflection Outlook

In essence, the primary aim of this thesis was to expand and enable the use of spectral data for pasture management, establishing useful links between the techniques of field spectroscopy and optimal imaging systems deployed at low-level flight. Rigorous analysis methods were demonstrated so that optimized sensor design can be expanded for many biophysical and biochemical attributes. This thesis has provided significant contributions to this goal.

Although sensors with a better combination of spectral bands could be designed, the current limitations of UAV seem to reside in the low radiometric quality of the end-products, the necessary processing power for the generation orthomosaic, and the consequent, poor generalization abilities of spectral models from one year to the next. As *"information consumers"*, farmers are not commited to an observational platform and will likely adopt a techological solution that offers an optimal trade-off between cost, autonomy and accuracy.

From our results, it seems plausible to hypothesize that a multispectral system with consistent radiometric response could be an effective tool for pasture management. From a

data science perspective, it seems that a spectral model, employing Cubist or an alternative regression-tree based algorithm, with periodic updates of sample-spectra pairs could provide a useful predictive model with large scale application. These sample-spectra pairs could be provided by a small number of end-users, updating global or regional spectral libraries, whenever a spectral response was found to be a high-leverage point through an unsupervised learning algorithm, similarly to the chemometric calibrations for benchtop NIRS. However, it is worth noting that this thesis has not explored temporal correlations, other sources of prior information such as weather data, agronomic practices, or seasonal patterns of perennial ryegrass. These sources of information and distinct modeling techniques seem to be reachable and adequate for precision pasture management.

As an Outlook, it seems reasonable to hypothesize that the tasks or orthomosaic generation and radiometric calibration will necessarily become a "commodity" so a number of industries could be developed exploring the potential of UAV spectral data and digital agriculture at the farm-scale. Currently, open-source photogrammetry tool-kits are widely available through initiatives such as OpenDroneMaps. However, a standard protocol for mission planning, radiometric calibration, and data processing is not in place for UAVs, constraining the adoption by non-scientific end-users. Consequently, it is safe to state that most users of UAV technology are still in either academia or research.

Nevertheless, the use of spectral data and imagery, regardless of its origin, is a new benchmark in digital agriculture. In this context, CubeSats and the new generation of satellites such as Sentinel-2, CBERS-4, or PRISMA may prove themselves to be a suitable alternative. Furthermore, radiometric calibration as a service and as global network (Flare$^{TM}$ - Labsphere) can enable accurate vicarious calibration (Schiller and Silny, 2010) for all these systems. Consequently, such would allow for a higher level of synergy and compatability between different constelations, leading to seamless data fusion and substantially increasing the frequency of revisit to a specific location.

Despite the upcoming synergies between different satellite constelations, all optical sensors are constrained by the presence of clouds, providing a competitive advantage to UAVs. If completely autonomous and low-cost, UAVs may become a usefull tool if able to operate under changing illumination conditions, provided that the operational pipeline is quick, reliable, and automated. It also seems plausible to suggest that the integration of UAVs and satellite data, alongside weather data, modeling strategies such as time-series, could provide better accuracies than what this thesis was able to achieve.

# References


Aasen, H., E. Honkavaara, A. Lucieer, and P. Zarco-Tejada (2018). "Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures, and Data Correction Workflows". *Remote Sensing* 10.7, 1091. DOI: 10.3390/rs10071091. URL: http://www.mdpi.com/2072-4292/10/7/1091.

Agisoft (2020). *Agisoft Metashape User Manual*. URL: https://www.agisoft.com/pdf/metashape-pro%7B%5C_%7D1%7B%5C_%7D5%7B%5C_%7Den.pdf (visited on 2020).

Alckmin, G. T., L. Kooistra, A. Lucieer, and R. Rawnsley (2019). "Feature Filtering and Selection for Dry Matter Estimation on Perennial Ryegrass: A Case Study Of Vegetation Indices". In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2/W13. 2/W13, 1827–1831. DOI: 10.5194/isprs-archives-XLII-2-W13-1827-2019. URL: https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W13/1827/2019/.

Ali, I., F. Cawkwell, E. Dwyer, B. Barrett, and S. Green (2016). "Satellite remote sensing of grasslands: from observation to management". *Journal of Plant Ecology* 9.6, 649–671. DOI: 10.1093/jpe/rtw005. URL: http://jpe.oxfordjournals.org/lookup/doi/10.1093/jpe/rtw005%20https://academic.oup.com/jpe/article-lookup/doi/10.1093/jpe/rtw005.

Allen, V., C. Batello, E. Berretta, J. Hodgson, M. Kothmann, X. Li, J. McIvor, J. Milne, C. Morris, A. Peeters, and M. Sanderson (2011). "An international terminology for grazing lands and grazing animals". *Grass and Forage Science* 66.1, 2–28. DOI: 10.1111/j.1365-2494.2010.00780.x. URL: http://doi.wiley.com/10.1111/j.1365-2494.2010.00780.x%20https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2494.2010.00780.x.

Anderegg, J., K. Yu, H. Aasen, A. Walter, F. Liebisch, and A. Hund (2020). "Spectral Vegetation Indices to Track Senescence Dynamics in Diverse Wheat Germplasm". *Frontiers in Plant Science* 10.January, 1–20. DOI: 10.3389/fpls.2019.01749. URL: https://www.frontiersin.org/article/10.3389/fpls.2019.01749/full.

Asner, G. P. (1998). "Biophysical and Biochemical Sources of Variability in Canopy Reflectance". *Remote Sensing of Environment* 64.3, 234–253. DOI: 10.1016/S0034-

4257(98)00014-5. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425798000145.

Australian Fodder Industry Association Standards (2013). *Fodders Trading Standards*.

Bajcsy, P. and P. Groves (2004). "Methodology for hyperspectral band selection". *Photogrammetric Engineering and Remote Sensing* 70.7, 793–802. DOI: 10.14358/PERS.70.7.793.

Baranoski, G. V. G. and D. Eng (2007). "An Investigation on Sieve and Detour Effects Affecting the Interaction of Collimated and Diffuse Infrared Radiation (750 to 2500 nm) With Plant Leaves". *IEEE Transactions on Geoscience and Remote Sensing* 45.8, 2593–2599. DOI: 10.1109/TGRS.2007.897427. URL: http://ieeexplore.ieee.org/document/4276874/.

Baret, F. and T. Fourty (1997). "Radiometric Estimates of Nitrogen Status of Leaves and Canopies". In: *Diagnosis of the Nitrogen Status in Crops*. Berlin, Heidelberg: Springer Berlin Heidelberg, 201–227. DOI: 10.1007/978-3-642-60684-7_12. URL: http://link.springer.com/10.1007/978-3-642-60684-7%7B%5C_%7D12.

Baret, F. and G. Guyot (1991). "Potentials and limits of vegetation indices for LAI and APAR assessment". *Remote Sensing of Environment* 35.2-3, 161–173. DOI: 10.1016/0034-4257(91)90009-U. URL: https://linkinghub.elsevier.com/retrieve/pii/003442579190009U.

Baret, F., V. Houles, and M. Guerif (2006). "Quantification of plant stress using remote sensing observations and crop models: the case of nitrogen management". *Journal of Experimental Botany* 58.4, 869–880. DOI: 10.1093/jxb/erl231. URL: https://academic.oup.com/jxb/article-lookup/doi/10.1093/jxb/erl231.

Baret, F. and B. Samuel (2008). *Advances in Land Remote Sensing*. Ed. by S. Liang. November 2014. Dordrecht: Springer Netherlands, 173–201. DOI: 10.1007/978-1-4020-6450-0. URL: http://dx.doi.org/10.1007/978-1-4020-6450-0%7B%5C_%7D12%7B%5C%%7D5Cnhttp://www.springerlink.com/index/10.1007/978-1-4020-6450-0%20http://link.springer.com/10.1007/978-1-4020-6450-0.

Baret, F. and S. Buis (2008). "Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems". *Advances in Land Remote Sensing: System, Modeling, Inversion and Application*, 173–201. DOI: 10.1007/978-1-4020-6450-0_7.

Barker, J. B., W. E. Woldt, B. D. Wardlow, C. M. Neale, M. S. Maguire, B. C. Leavitt, and D. M. Heeren (2020). "Calibration of a common shortwave multispectral camera system for quantitative agricultural applications". *Precision Agriculture* 21.4, 922–935. DOI: 10.1007/s11119-019-09701-6. URL: https://doi.org/10.1007/s11119-019-09701-6.

Barnes, E. M., T. R. Clarke, S. E. Richards, P. D. Colaizzi, J. Haberland, M. Kostrzewski, P. Waller, C. Choi, E. Riley, T. Thompson, R. J. Lascano, H. Li, and M. S. Moran (2000).

"Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data." In: *Proceedings of the 5th International Conference on Precision Agriculture*. Ed. by P. C. Robert, R. H. Rust, and W. E. Larson. Bloomington, MN: American Society of Agronomy, 1–15.

Beer (1852). "Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten". *Annalen der Physik und Chemie* 162.5, 78–88. DOI: 10.1002/andp.18521620505. URL: http://doi.wiley.com/10.1002/andp.18521620505.

Beldman, A., J. Reijs, C. Daatselaar, and G. Doornewaard (2020). *De Nederlandse melkveehouderij in 2030 : verkenning van mogelijke ontwikkelingen op basis van economische modellering*. Tech. rep. DOI: 10.18174/532156. URL: https://research.wur.nl/en/publications/4153ee39-6686-4d89-9eff-5d67962669d3.

Berger, K., J. Verrelst, J.-B. Féret, T. Hank, M. Wocher, W. Mauser, and G. Camps-Valls (2020). "Retrieval of aboveground crop nitrogen content with a hybrid machine learning method". *International Journal of Applied Earth Observation and Geoinformation* 92.June, 102174. DOI: 10.1016/j.jag.2020.102174. URL: https://doi.org/10.1016/j.jag.2020.102174%20https://linkinghub.elsevier.com/retrieve/pii/S0303243420303500.

Biewer, S., T. Fricke, and M. Wachendorf (2009). "Development of Canopy Reflectance Models to Predict Forage Quality of Legume-Grass Mixtures". *Crop Science* 49.5, 1917–1926. DOI: 10.2135/cropsci2008.11.0653. URL: http://doi.wiley.com/10.2135/cropsci2008.11.0653.

Bonham-Carter, G. (1988). "Numerical procedures and computer program for fitting an inverted gaussian model to vegetation reflectance data". *Computers & Geosciences* 14.3, 339–356. DOI: 10.1016/0098-3004(88)90065-9. URL: https://linkinghub.elsevier.com/retrieve/pii/0098300488900659.

Breiman, L. (1984). "Random Forests". *Machine Learning* 45.1, 5–32. DOI: https://doi.org/10.1023/A:1010933404324.

– (1994). "Bagging Predictors". 421.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone (2017). *Classification And Regression Trees*. The Wadsworth and Brooks-Cole statistics-probability series November. Boca Raton: CRC Press, 1–358. DOI: 10.1201/9781315139470. URL: https://www.taylorfrancis.com/books/9781351460491.

Burkart, A., S. Cogliati, A. Schickling, and U. Rascher (2014). "A Novel UAV-Based Ultra-Light Weight Spectrometer for Field Spectroscopy". *IEEE Sensors Journal* 14.1, 62–67. DOI: 10.1109/JSEN.2013.2279720. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6587080%20http://ieeexplore.ieee.org/document/6587080/.

Cao, D.-S., Z.-K. Deng, M.-F. Zhu, Z.-J. Yao, J. Dong, and R.-G. Zhao (2017). "Ensemble partial least squares regression for descriptor selection, outlier detection, applicability

domain assessment, and ensemble modeling in QSAR/QSPR modeling". *Journal of Chemometrics* 31.11, e2922. DOI: 10.1002/cem.2922. URL: http://doi.wiley.com/10.1002/cem.2922.

Capolupo, A., L. Kooistra, C. Berendonk, L. Boccia, and J. Suomalainen (2015). "Estimating Plant Traits of Grasslands from UAV-Acquired Hyperspectral Images: A Comparison of Statistical Approaches". *ISPRS International Journal of Geo-Information* 4.4, 2792–2820. DOI: 10.3390/ijgi4042792. URL: http://www.mdpi.com/2220-9964/4/4/2792.

Carlson, T. N. and D. A. Ripley (1997). "On the relation between NDVI, fractional vegetation cover, and leaf area index". *Remote Sensing of Environment* 62.3, 241–252. DOI: 10.1016/S0034-4257(97)00104-1. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425797001041.

Caruana, R., A. Niculescu-Mizil, G. Crew, and A. Ksikes (2004). "Ensemble selection from libraries of models". In: *Twenty-first international conference on Machine learning - ICML '04*. New York, New York, USA: ACM Press, 18. DOI: 10.1145/1015330.1015432. URL: http://portal.acm.org/citation.cfm?doid=1015330.1015432.

Carvalho, D. V., E. M. Pereira, and J. S. Cardoso (2019). "Machine Learning Interpretability: A Survey on Methods and Metrics". *Electronics* 8.8, 832. DOI: 10.3390/electronics8080832. URL: https://www.mdpi.com/2079-9292/8/8/832.

Chapman, D. F., J. Tharmaraj, M. Agnusdei, and J. Hill (2012). "Regrowth dynamics and grazing decision rules: further analysis for dairy production systems based on perennial ryegrass (Lolium perenne L.) pastures". *Grass and Forage Science* 67.1, 77–95. DOI: 10.1111/j.1365-2494.2011.00824.x. URL: http://doi.wiley.com/10.1111/j.1365-2494.2011.00824.x.

Cho, M. A., A. Skidmore, F. Corsi, S. E. van Wieren, and I. Sobhan (2007). "Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression". en. *International Journal of Applied Earth Observation and Geoinformation* 9.4, 414–424. DOI: 10.1016/j.jag.2007.02.001. URL: http://linkinghub.elsevier.com/retrieve/pii/S030324340700013X%20https://linkinghub.elsevier.com/retrieve/pii/S030324340700013X.

Christie, K. M., A. P. Smith, R. P. Rawnsley, M. T. Harrison, and R. J. Eckard (2018). "Simulated seasonal responses of grazed dairy pastures to nitrogen fertilizer in SE Australia: Pasture production". *Agricultural Systems* 166, 36–47. DOI: 10.1016/j.agsy.2018.07.010. URL: https://linkinghub.elsevier.com/retrieve/pii/S0308521X1731051X.

Clevers, J., G. van der Heijden, S. Verzakov, and M. Schaepman (2007). "Estimating Grassland Biomass Using SVM Band Shaving of Hyperspectral Data". *Photogrammetric Engineering & Remote Sensing* 73.10, 1141–1148. DOI: 10.14358/PERS.73.10.1141. URL: http://skgr0103.wur.nl/%7B~%7Dschae001/paper/2007%7B%5C_

%7DBandShaving%7B%5C_%7DPERS%7B%5C_%7DJC.pdf%20http://openurl.ingenta.com/content/xref?genre=article%7B%5C&%7Dissn=0099-1112%7B%5C&%7Dvolume=73%7B%5C&%7Dissue=10%7B%5C&%7Dspage=1141.

Collins, W. (1978). "Remote Sensing of Crop Type and Maturity." *Photogrammetric Engineering and Remote Sensing* 44.1, 42–55.

Cortes, C. and V. Vapnik (1995). "Support-Vector Networks". *Machine Learning* 20.3. Ed. by L. Saitta, 273–297. DOI: 10.1023/A:1022627411411. URL: https://doi.org/10.1007/BF00994018%20http://link.springer.com/10.1007/BF00994018.

Curran, P. J. (1989). "Remote sensing of foliar chemistry". *Remote Sensing of Environment* 30.3, 271–278. DOI: 10.1016/0034-4257(89)90069-2. URL: http://www.sciencedirect.com/science/article/pii/0034425789900692%20https://linkinghub.elsevier.com/retrieve/pii/0034425789900692.

– (1994). "Imaging spectrometry". *Progress in Physical Geography: Earth and Environment* 18.2, 247–266. DOI: 10.1177/030913339401800204. URL: http://journals.sagepub.com/doi/10.1177/030913339401800204.

Dairy Australia (2020). *Annual Report*. Tech. rep. Southbank: Dairy Australia. URL: https://www.rtda.gov.rw/fileadmin/templates/documents/Annual%7B%5C_%7DReport%7B%5C_%7D2017%7B%5C_%7D2018%7B%5C_%7DFINAL.pdf;.

Darvishzadeh, R., A. Skidmore, M. Schlerf, C. Atzberger, F. Corsi, and M. Cho (2008). "LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements". *ISPRS Journal of Photogrammetry and Remote Sensing* 63.4, 409–426. DOI: 10.1016/j.isprsjprs.2008.01.001. URL: https://linkinghub.elsevier.com/retrieve/pii/S0924271608000166.

Daughtry, C. (2000). "Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance". *Remote Sensing of Environment* 74.2, 229–239. DOI: 10.1016/S0034-4257(00)00113-9. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425700001139.

Dawson, T. P. and P. J. Curran (1998). "Technical note A new technique for interpolating the reflectance red edge position". *International Journal of Remote Sensing* 19.11, 2133–2139. DOI: 10.1080/014311698214910.

De Rosa, D., B. Basso, M. Fasiolo, J. Friedl, B. Fulkerson, P. R. Grace, and D. W. Rowlings (2021). "Predicting pasture biomass using a statistical model and machine learning algorithm implemented with remotely sensed imagery". *Computers and Electronics in Agriculture* 180.October 2020, 105880. DOI: 10.1016/j.compag.2020.105880. URL: https://doi.org/10.1016/j.compag.2020.105880%20https://linkinghub.elsevier.com/retrieve/pii/S0168169920330854.

Delegido, J., J. Verrelst, L. Alonso, and J. Moreno (2011). "Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content". *Sensors*

11.7, 7063–7081. DOI: 10.3390/s110707063. URL: http://www.mdpi.com/1424-8220/11/7/7063.

Dinno, A. (2015). "Nonparametric pairwise multiple comparisons in independent groups using Dunn's test". *Stata Journal* 15.1, 292–300. DOI: 10.1177/1536867x1501500117.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. Mcclean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, and S. Lautenbach (2013). "Collinearity: A review of methods to deal with it and a simulation study evaluating their performance". *Ecography* 36.1, 27–46. DOI: 10.1111/j.1600-0587.2012.07348.x.

Dubbini, M., A. Pezzuolo, M. De Giglio, M. Gattelli, L. Curzio, D. Covi, T. Yezekyan, and F. Marinello (2017). "Last generation instrument for agriculture multispectral data collection". *Agricultural Engineering International: CIGR Journal* 19.1, 87–93.

Earle, D. and A. McGowan (1979). "Evaluation and calibration of an automated rising plate meter for estimating dry matter yield of pasture". *Australian Journal of Experimental Agriculture* 19.98, 337. DOI: 10.1071/EA9790337. URL: http://www.publish.csiro.au/?paper=EA9790337.

Edirisinghe, A., D. Clark, and D. Waugh (2012). "Spatio-temporal modelling of biomass of intensively grazed perennial dairy pastures using multispectral remote sensing". *International Journal of Applied Earth Observation and Geoinformation* 16.1, 5–16. DOI: 10.1016/j.jag.2011.11.006. URL: http://dx.doi.org/10.1016/j.jag.2011.11.006%20https://linkinghub.elsevier.com/retrieve/pii/S0303243411001711.

Esbensen, K. and P. Geladi (1990). "The start and early history of chemometrics: Selected interviews. Part 2". *Journal of Chemometrics* 4.6, 389–412. DOI: 10.1002/cem.1180040604.

Fahey, G. C., J. S. Shenk, and M. O. Westerhaus (1994). "The Application of near Infrared Reflectance Spectroscopy (NIRS) to Forage Analysis". 1978, 406–449. DOI: 10.2134/1994.foragequality.c10.

Fallet, C. and L. M. Domenzain (2018). "Necessary steps for the systematic calibration of a multispectral imaging system to achieve a targetless workflow in reflectance estimation: a study of Parrot SEQUOIA for precision agriculture". In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV*. Ed. by D. W. Messinger and M. Velez-Reyes. Vol. 10644. May 2018. International Society for Optics and Photonics. SPIE, 42. DOI: 10.1117/12.2304334. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10644/2304334/Necessary-steps-for-the-systematic-calibration-of-a-multispectral-imaging/10.1117/12.2304334.full%20https://doi.org/10.1117/12.2304334.

Fariña, S. R., S. C. Garcia, W. J. Fulkerson, and I. M. Barchia (2011). "Pasture-based dairy farm systems increasing milk production through stocking rate or milk yield per cow: pasture and animal responses". *Grass and Forage Science* 66.3, 316–332. DOI:

10.1111/j.1365-2494.2011.00795.x. URL: http://doi.wiley.com/10.1111/j.1365-2494.2011.00795.x.

Fernández-Delgado, M., M. Sirsat, E. Cernadas, S. Alawadi, S. Barro, and M. Febrero-Bande (2019). "An extensive experimental survey of regression methods". *Neural Networks* 111, 11–34. DOI: 10.1016/j.neunet.2018.12.010. URL: https://doi.org/10.1016/j.neunet.2018.12.010%20https://linkinghub.elsevier.com/retrieve/pii/S0893608018303411.

Filella, I. and J. Peñuelas (1994). "The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status". *International Journal of Remote Sensing* 15.7, 1459–1470. DOI: 10.1080/01431169408954177.

Franzini, M., G. Ronchetti, G. Sona, and V. Casella (2019). "Geometric and Radiometric Consistency of Parrot Sequoia Multispectral Imagery for Precision Agriculture Applications". *Applied Sciences* 9.24, 5314. DOI: 10.3390/app9245314. URL: https://www.mdpi.com/2076-3417/9/24/5314.

Fraser, B. T. and R. G. Congalton (2018). "Issues in Unmanned Aerial Systems (UAS) data collection of complex forest environments". *Remote Sensing* 10.6. DOI: 10.3390/rs10060908.

Friedman, J. H. (1991). "Multivariate Adaptive Regression Splines". *The Annals of Statistics* 19.1, 1–67. DOI: 10.1214/aos/1176347963. URL: https://doi.org/10.1214/aos/1176347963%20https://www.jstor.org/stable/2241837.

Fulkerson, W. J., J. S. Neal, C. F. Clark, A. Horadagoda, K. S. Nandra, and I. Barchia (2007). "Nutritive value of forage species grown in the warm temperate climate of Australia for dairy cows: Grasses and legumes". *Livestock Science* 107.2-3, 253–264. DOI: 10.1016/j.livsci.2006.09.029.

Fussell, J., D. Rundquist, and J. A. Harrington (1986). "On defining remote sensing." *Photogrammetric Engineering & Remote Sensing* 52.9, 1507–1511.

Gao, B.-C. and A. F. Goetz (1994). "Extraction of dry leaf spectral features from reflectance spectra of green vegetation". *Remote Sensing of Environment* 47.3, 369–374. DOI: 10.1016/0034-4257(94)90104-X. URL: https://linkinghub.elsevier.com/retrieve/pii/003442579490104X.

García, S. C., M. R. Islam, C. E. F. Clark, and P. M. Martin (2014). "Kikuyu-based pasture for dairy production: a review". *Crop and Pasture Science* 65.8, 787. DOI: 10.1071/CP13414. URL: http://www.publish.csiro.au/?paper=CP13414.

Gargiulo, J. I., C. R. Eastwood, S. C. Garcia, and N. A. Lyons (2018). "Dairy farmers with larger herd sizes adopt more precision dairy technologies". *Journal of Dairy Science* 101.6, 5466–5473. DOI: 10.3168/jds.2017-13324.

Geipel, J. and A. Korsaeth (2017). "Hyperspectral Aerial Imaging for Grassland Yield Estimation". *Advances in Animal Biosciences* 8.02, 770–775. DOI: 10.1017/

S2040470017000619. URL: https://www.cambridge.org/core/product/identifier/S2040470017000619/type/journal%7B%5C_%7Darticle.

Gevaert, C. M., J. Suomalainen, J. Tang, and L. Kooistra (2015). "Generation of Spectral–Temporal Response Surfaces by Combining Multispectral Satellite and Hyperspectral UAV Imagery for Precision Agriculture Applications". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8.6, 3140–3146. DOI: 10.1109/JSTARS.2015.2406339. URL: http://ieeexplore.ieee.org/document/7058421/.

Gianelle, D. and F. Guastella (2007). "Nadir and off-nadir hyperspectral field data: Strengths and limitations in estimating grassland biophysical characteristics". *International Journal of Remote Sensing* 28.7, 1547–1560. DOI: 10.1080/01431160600658180.

Gilliot, J.-m., L. M. Domenzain, R. Faroux, and J. Michelin (2018). "Correction of in-flight luminosity variations in multispectral UAS images, using a luminosity sensor and camera pair for improved biomass estimation in precision agriculture". In: *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping III*. Ed. by J. A. Thomasson, M. McKee, and R. J. Moorhead. Vol. 1066405. May 2018. Orlando: SPIE, 4. DOI: 10.1117/12.2303804. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10664/2303804/Correction-of-in-flight-luminosity-variations-in-multispectral-UAS-images/10.1117/12.2303804.full.

Gitelson, A. and M. N. Merzlyak (1994). "Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation". *Journal of Plant Physiology* 143.3, 286–292. DOI: 10.1016/S0176-1617(11)81633-0. URL: https://linkinghub.elsevier.com/retrieve/pii/S0176161711816330.

Gitelson, A. A. (2004). "Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation". *Journal of Plant Physiology* 161.2, 165–173. DOI: 10.1078/0176-1617-01176.

Gitelson, A. A., Y. J. Kaufman, and M. N. Merzlyak (1996). "Use of a green channel in remote sensing of global vegetation from EOS-MODIS". *Remote Sensing of Environment* 58.3, 289–298. DOI: 10.1016/S0034-4257(96)00072-7. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425796000727.

Guerini Filho, M., T. M. Kuplich, and F. L. F. D. Quadros (2020). "Estimating natural grassland biomass by vegetation indices using Sentinel 2 remote sensing data". *International Journal of Remote Sensing* 41.8, 2861–2876. DOI: 10.1080/01431161.2019.1697004. URL: https://doi.org/10.1080/01431161.2019.1697004%20https://www.tandfonline.com/doi/full/10.1080/01431161.2019.1697004.

Guyot, G. and F. Baret (1988). "Utilisation de la Haute Resolution Spectrale pour Suivre L etat des Couverts Vegetaux". In: *Spectral Signatures of Objects in Remote Sensing*. Ed. by T. Guyenne and J. Hunt. Vol. 287. ESA Special Publication March 1988. Aussois: European Space Agency, 279. DOI: 10.1007/s13398-014-0173-7.2. arXiv: 9809069v1 [arXiv:gr-qc].

Haas, R. H., D. W. Deering, J. W. Rouse, and J. A. Schell (1975). "Monitoring vegetation conditions from Landsat for use in range management". In: *NASA Earth Resources Survey Symposium Proc., Houston, Texas*. Vol. 1, 43–52.

Hadamard, J. (1902). "Sur les problems aux derivees patielles et leur signification physique". *Princeton University Bulletin* 13, 49–52. URL: https://ci.nii.ac.jp/naid/10030001312/en/.

Hakala, T., E. Honkavaara, H. Saari, J. Mäkynen, J. Kaivosoja, L. Pesonen, and I. Pölönen (2013). "Spectral Imaging From UAVs Under Varying Illumination Conditions". *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-1/W2, 189–194. DOI: 10.5194/isprsarchives-XL-1-W2-189-2013. URL: http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-1-W2/189/2013/.

Hall, A., L. Turner, L. Irvine, and S. Kilpatrick (2017). "Pasture management and extension on Tasmanian dairy farms - who measures up?" *Rural Extension and Innovation Systems Journal* 13.2, 32–40.

Hall, A., L. Turner, and S. Kilpatrick (2019). "Understanding Tasmanian dairy farmer adoption of pasture management practices: A Theory of Planned Behaviour approach". *Animal Production Science* 59.10, 1941–1950. DOI: 10.1071/AN18321.

Hanrahan, L., N. McHugh, T. Hennessy, B. Moran, R. Kearney, M. Wallace, and L. Shalloo (2018). "Factors associated with profitability in pasture-based systems of milk production". *Journal of Dairy Science* 101.6, 5474–5485. DOI: 10.3168/jds.2017-13223.

Hansen, P. M. and J. K. Schjoerring (2003). "Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression". *Remote Sensing of Environment* 86.4, 542–553. DOI: 10.1016/S0034-4257(03)00131-7.

Hennessy, D., L. Delaby, A. van den Pol-van Dasselaar, and L. Shalloo (2020). "Increasing grazing in dairy cow milk production systems in Europe". *Sustainability (Switzerland)* 12.6, 1–15. DOI: 10.3390/su12062443.

Horler, D. N., M. Dockray, and J. Barber (1983). "The red edge of plant leaf reflectance". *International Journal of Remote Sensing* 4.2, 273–288. DOI: 10.1080/01431168308948546.

Hotelling, H. (1931). "The Generalization of Student's Ratio". *The Annals of Mathematical Statistics* 2.3, 360–378. DOI: 10.1214/aoms/1177732979. URL: https://doi.org/10.

1214/aoms/1177732979%20http://projecteuclid.org/euclid.aoms/1177732979%20http://link.springer.com/10.1007/978-1-4612-0919-5%7B%5C_%7D4.

Hoving, I. E., D. A. J. Starmans, J. A. Booij, I. Kuiper, and G. Holshof (2018). "Amazing grazing: grass growth measurements with remote sensing techniques." In: ed. by B. Horan, D. Hennessy, M. O'Donovan, E. Kennedy, B. McCarthy, J. A. Finn, and B. O'Brien. Fermoy, Irish Republic: Teagasc, Animal & Grassland Research and Innovation Centre, 860–862.

Imran, H. A., D. Gianelle, D. Rocchini, M. Dalponte, M. P. Martín, K. Sakowska, G. Wohlfahrt, and L. Vescovo (2020). "VIS-NIR, red-edge and NIR-shoulder based normalized vegetation indices response to co-varying leaf and Canopy structural traits in heterogeneous grasslands". *Remote Sensing* 12.14, 2254. DOI: 10.3390/rs12142254.

Jaccard, P. (1901). "Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines". *Bulletin de la Société Vaudoise des Sciences Naturelles* 37.January 1901, 241–272. DOI: 10.5169/seals-266440.

Jago, J., C. Eastwood, K. Kerrisk, and I. Yule (2013). "Precision dairy farming in Australasia: Adoption, risks and opportunities". *Animal Production Science* 53.9, 907–916. DOI: 10.1071/AN12330.

Jiang, J., W. Cai, H. Zheng, T. Cheng, Y. Tian, Y. Zhu, R. Ehsani, Y. Hu, Q. Niu, L. Gui, and X. Yao (2019). "Using digital cameras on an unmanned aerial vehicle to derive optimum color vegetation indices for leaf nitrogen concentration monitoring in winter wheat". *Remote Sensing* 11.22. DOI: 10.3390/rs11222667.

Joint Committee for Guides in Metrology (2008). "Evaluation of measurement data — Guide to the expression of uncertainty in measurement". URL: http://www.bipm.org/en/publications/guides/gum.html.

Jones, G., N. S. Wade, J. P. Baker, and E. M. Ranck (1987). "Use of Near Infrared Reflectance Spectroscopy in Forage Testing". *Journal of Dairy Science* 70.5, 1086–1091. DOI: 10.3168/jds.S0022-0302(87)80115-7. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022030287801157.

Jongschaap, R. E. and R. Booij (2004). "Spectral measurements at different spatial scales in potato: Relating leaf, plant and canopy nitrogen status". *International Journal of Applied Earth Observation and Geoinformation* 5.3, 205–218. DOI: 10.1016/j.jag.2004.03.002.

Josse, J. and F. Husson (2016). "missMDA : A Package for Handling Missing Values in Multivariate Data Analysis". *Journal of Statistical Software* 70.1. DOI: 10.18637/jss.v070.i01. URL: http://www.jstatsoft.org/v70/i01/.

Kallenbach, R. L. (2015). "Describing the dynamic: Measuring and assessing the value of plants in the pasture". *Crop Science* 55.6, 2531–2539. DOI: 10.2135/cropsci2015.01.0065.

Karunaratne, S., E. Morse-mcnabb, A. Thomson, D. Stayches, and J. Jacobs (2019). "Paddock scale modelling and mapping of dry matter yield using UAV derived datasets : A case from dairy farming systems in Victoria". In: *Proceedings of the 2019 Agronomy Australia Conference*. Wagga Wagga, 6–9.

Karunaratne, S., A. Thomson, E. Morse-McNabb, J. Wijesingha, D. Stayches, A. Copland, and J. Jacobs (2020). "The Fusion of Spectral and Structural Datasets Derived from an Airborne Multispectral Sensor for Estimation of Pasture Dry Matter Yield at Paddock Scale with Time". *Remote Sensing* 12.12, 2017. DOI: 10.3390/rs12122017. URL: https://www.mdpi.com/2072-4292/12/12/2017.

Kattenborn, T., F. Schiefer, P. Zarco-Tejada, and S. Schmidtlein (2019). "Advantages of retrieving pigment content [$\mu$g/cm2] versus concentration [%] from canopy reflectance". *Remote Sensing of Environment* 230.September 2018, 111195. DOI: 10.1016/j.rse.2019.05.014. URL: https://doi.org/10.1016/j.rse.2019.05.014.

Kawamura, K., K. Betteridge, I. D. Sanches, M. P. Tuohy, D. E. Costall, and Y. Inoue (2009). "Field radiometer with canopy pasture probe as a potential tool to estimate and map pasture biomass and mineral components: A case study in the Lake Taupo catchment, New Zealand". *New Zealand Journal of Agricultural Research* 52.4, 417–434. DOI: 10.1080/00288230909510524.

Kawamura, K., N. Watanabe, S. Sakanoue, and Y. Inoue (2008). "Estimating forage biomass and quality in a mixed sown pasture based on partial least squares regression with waveband selection". *Grassland Science* 54.3, 131–145. DOI: 10.1111/j.1744-697x.2008.00116.x.

Kawamura, K., N. Watanabe, S. Sakanoue, H.-J. Lee, Y. Inoue, and S. Odagawa (2010). "Testing genetic algorithm as a tool to select relevant wavebands from field hyperspectral data for estimating pasture mass and quality in a mixed sown pasture using partial least squares regression". *Grassland Science* 56.4, 205–216. DOI: 10.1111/j.1744-697X.2010.00196.x. URL: http://doi.wiley.com/10.1111/j.1744-697X.2010.00196.x.

Kelcey, J. and A. Lucieer (2012). "Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing". en. *Remote Sensing* 4.5, 1462–1493. DOI: 10.3390/rs4051462. URL: http://www.mdpi.com/2072-4292/4/5/1462/.

Kim, H., W.-Y. Loh, Y.-S. Shih, and P. Chaudhuri (2007). "Visualizable and interpretable regression models with good prediction power". *IIE Transactions* 39.6, 565–579. DOI: 10.1080/07408170600897502. URL: https://www.tandfonline.com/doi/full/10.1080/07408170600897502.

Kjeldahl, K. and R. Bro (2010). "Some common misunderstandings in chemometrics". *Journal of Chemometrics* 24.7-8, 558–564. DOI: 10.1002/cem.1346. URL: http://doi.wiley.com/10.1002/cem.1346.

Knaus, W. (2016). "Perspectives on pasture versus indoor feeding of dairy cows". *Journal of the Science of Food and Agriculture* 96.1, 9–17. DOI: 10.1002/jsfa.7273.

Kokaly, R. (1999). "Spectroscopic Determination of Leaf Biochemistry Using Band-Depth Analysis of Absorption Features and Stepwise Multiple Linear Regression". *Remote Sensing of Environment* 67.3, 267–287. DOI: 10.1016/S0034-4257(98)00084-4. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425798000844.

Korte, C. J., B. R. Watkin, and W. Harris (1988). "Use of residual leaf area index and light interception as criteria for spring-grazing management of a ryegrass-dominant pasture". *New Zealand Journal of Agricultural Research* 25.3, 309–319. DOI: 10.1080/00288233.1982.10417892.

Kuhn, M. (2008). "Building Predictive Models in R Using the caret Package". *Journal of Statistical Software* 28.5, 1–26. DOI: 10.18637/jss.v028.i05. URL: https://www.jstatsoft.org/v028/i05%20http://www.jstatsoft.org/v28/i05/.

– (2016). *Desirability: Function Optimization and Ranking via Desirability Functions*. URL: https://cran.r-project.org/package=desirability.

Kuhn, M. and K. Johnson (2013). *Applied Predictive Modeling*. Vol. 26. New York: Springer, 620. DOI: 10.1007/978-1-4614-6849-3. URL: http://www.amazon.com/Applied-Predictive-Modeling-Max-Kuhn/dp/1461468485/ref=pd%7B%5C_%7Dbxgy%7B%5C_%7Db%7B%5C_%7Dimg%7B%5C_%7Dz.

– (2019). *Feature Engineering and Selection : a Practical Approach for Predictive Models*. 1st Editio. Boca Raton: Chapman and Hall/CRC, 314. DOI: 10.4324/9781315108230. URL: http://www.feat.engineering/index.html.

Kuhn, M. and H. Wickham (2020). *Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles*. Boston, MA: CRAN. URL: https://www.tidymodels.org.

Lamb, D. W., M. Steyn-ross, P. Schaares, M. M. Hanna, W. Silvester, and A. Steyn-Ross (2002). "Estimating leaf nitrogen concentration in ryegrass (Lolium spp.) pasture using the chlorophyll red-edge: Theoretical modelling and experimental observations". *International Journal of Remote Sensing* 23.18, 3619–3648. DOI: 10.1080/01431160110114529.

Lantinga, E. A., M. Nassiri, and M. J. Kropff (1999). "Modelling and measuring vertical light absorption within grass–clover mixtures". *Agricultural and Forest Meteorology* 96.1-3, 71–83. DOI: 10.1016/S0168-1923(99)00040-4. URL: https://linkinghub.elsevier.com/retrieve/pii/S0168192399000404.

Latham, J., R. Cumani, I. Rosati, and M. Bloise (2014). *Global Land Cover SHARE (GLC-SHARE) database Beta-Release Version 1.0*. Tech. rep., 1–39. URL: http://www.fao.org/uploads/media/glc-share-doc.pdf.

Lehnert, L. W., H. Meyer, W. A. Obermeier, B. Silva, B. Regeling, and J. Bendix (2018). "Hyperspectral Data Analysis in R: the hsdar Package". *Journal of Statistical Software* 89.12. DOI: 10.18637/jss.v089.i12. arXiv: 1805.05090. URL: http://arxiv.

org/abs/1805.05090%20http://dx.doi.org/10.18637/jss.v089.i12%20http://www.jstatsoft.org/v89/i12/.

Lichtenthaler, H. K., M. Lang, M. Sowinska, F. Heisel, and J. A. Miehé (1996). "Detection of vegetation stress via a new high resolution fluorescence imaging system". *Journal of Plant Physiology* 148.5, 599–612. DOI: 10.1016/S0176-1617(96)80081-2.

Lin, L. I.-K. (1989). "A Concordance Correlation Coefficient to Evaluate Reproducibility". *Biometrics* 45.1, 255. DOI: 10.2307/2532051. URL: https://www.jstor.org/stable/2532051?origin=crossref.

Lorenz, H., T. Reinsch, S. Hess, and F. Taube (2019). "Is low-input dairy farming more climate friendly? A meta-analysis of the carbon footprints of different production systems". *Journal of Cleaner Production* 211, 161–170. DOI: 10.1016/j.jclepro.2018.11.113. URL: https://doi.org/10.1016/j.jclepro.2018.11.113.

Lu, H., T. Fan, P. Ghimire, and L. Deng (2020). "Experimental evaluation and consistency comparison of UAV multispectral minisensors". *Remote Sensing* 12.16, 1–19. DOI: 10.3390/RS12162542.

Mac Arthur, A. and I. Robinson (2015). "A critique of field spectroscopy and the challenges and opportunities it presents for remote sensing for agriculture, ecosystems, and hydrology". In: *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVII*. Ed. by C. M. U. Neale and A. Maltese. Vol. 9637. International Society for Optics and Photonics. Toulouse: SPIE, 963705. DOI: 10.1117/12.2201046. URL: https://doi.org/10.1117/12.2201046%20http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2201046.

Mac Arthur, A. A., C. MacLellan, and T. J. Malthus (2007). "The implications of non-uniformity in fields-of-view of commonly used field spectroradiometers". In: *2007 IEEE International Geoscience and Remote Sensing Symposium*. January. Barcelona: IEEE, 2890–2893. DOI: 10.1109/IGARSS.2007.4423447. URL: http://ieeexplore.ieee.org/document/4423447/.

Machado, C. F., S. T. Morris, J. Hodgson, and M. Fathalla (2005). "Seasonal changes of herbage quality within a New Zealand beef cattle finishing pasture". *New Zealand Journal of Agricultural Research* 48.2, 265–270. DOI: 10.1080/00288233.2005.9513655.

Maes, W. H. and K. Steppe (2019). "Perspectives for Remote Sensing with Unmanned Aerial Vehicles in Precision Agriculture". *Trends in Plant Science* 24.2, 152–164. DOI: 10.1016/j.tplants.2018.11.007. URL: https://doi.org/10.1016/j.tplants.2018.11.007.

Maire, G. le, C. François, K. Soudani, D. Berveiller, J. Y. Pontailler, N. Bréda, H. Genet, H. Davi, and E. Dufrêne (2008). "Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass". *Remote Sensing of Environment* 112.10, 3846–3864. DOI: 10.1016/j.rse.2008.06.005.

Mamaghani and Salvaggio (2019). "Multispectral Sensor Calibration and Characterization for sUAS Remote Sensing". *Sensors* 19.20, 4453. DOI: 10.3390/s19204453. URL: https://www.mdpi.com/1424-8220/19/20/4453.

Marten, G., J. Shenk, and F. Barton (1989). *Near infrared reflectance spectroscopy (NIRS): analysis of forage quality*. Vol. 643. Washington, DC, USA, 1–110.

Massey, F. J. (1951). "The Kolmogorov-Smirnov Test for Goodness of Fit". *Journal of the American Statistical Association* 46.253, 68–78. DOI: 10.1080/01621459.1951.10500769. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1951.10500769.

Mayer, Z. A. and J. E. Knowles (2015). *caretEnsemble: Ensembles of caret models.* URL: http://cran.r-project.org/package=caretEnsemble.

McClure, W. F. (1993). "More on Derivatives: Part 1. Segments, Gaps and "Ghosts"". *NIR news* 4.6, 12–13. DOI: 10.1255/nirn.220.

McEvoy, M., M. O'Donovan, E. Kennedy, J. P. Murphy, L. Delaby, and T. M. Boland (2009). "Effect of pregrazing herbage mass and pasture allowance on the lactation performance of Holstein-Friesian dairy cows". *Journal of Dairy Science* 92.1, 414–422. DOI: 10.3168/jds.2008-1313. URL: http://dx.doi.org/10.3168/jds.2008-1313.

Mehmood, T., K. H. Liland, L. Snipen, and S. Sæbø (2012). "A review of variable selection methods in Partial Least Squares Regression". *Chemometrics and Intelligent Laboratory Systems* 118, 62–69. DOI: 10.1016/j.chemolab.2012.07.010. URL: http://dx.doi.org/10.1016/j.chemolab.2012.07.010%20https://linkinghub.elsevier.com/retrieve/pii/S0169743912001542.

Mevik, B. and R. Wehrens (2015). *Introduction to the pls Package*. URL: https://cran.r-project.org/web/packages/pls/vignettes/pls-manual.pdf.

Meyer, H. and E. Pebesma (2020). "Predicting Into Unknown Space? Estimating the Area of Applicability of Spatial Prediction Models". *arXiv*, 1–16. arXiv: 2005.07939.

Meyer, H., C. Reudenbach, T. Hengl, M. Katurji, and T. Nauss (2018). "Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation". *Environmental Modelling & Software* 101, 1–9. DOI: 10.1016/j.envsoft.2017.12.001. URL: https://doi.org/10.1016/j.envsoft.2017.12.001%20https://linkinghub.elsevier.com/retrieve/pii/S1364815217310976.

Meyer, H., C. Reudenbach, S. Wöllauer, and T. Nauss (2019). "Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction". *Ecological Modelling* 411. DOI: 10.1016/j.ecolmodel.2019.108815. arXiv: 1908.07805.

Michez, A., P. Lejeune, S. Bauwens, A. Herinaina, Y. Blaise, E. Castro Muñoz, F. Lebeau, J. Bindelle, A. A. Lalaina Herinaina, Y. Blaise, E. C. Muñoz, F. Lebeau, and J. Bindelle (2019). "Mapping and Monitoring of Biomass and Grazing in Pasture with an Unmanned

Aerial System". *Remote Sensing* 11.5, 473. DOI: 10.3390/rs11050473. URL: https://www.mdpi.com/2072-4292/11/5/473.

Michez, A., L. Philippe, K. David, C. Sébastien, D. Christian, and J. Bindelle (2020). "Can low-cost unmanned aerial systems describe the forage quality heterogeneity? Insight from a timothy pasture case study in southern Belgium". *Remote Sensing* 12.10. DOI: 10.3390/rs12101650.

Miller, E. J. (1935). "The basic amino-acids of typical forage grass proteins". *Biochemical Journal* 29.10, 2344–2350. DOI: 10.1042/bj0292344. URL: https://portlandpress.com/biochemj/article/29/10/2344/20251/The-basic-aminoacids-of-typical-forage-grass.

Milton, E. J. (1987). "Principles of field spectroscopy". *International Journal of Remote Sensing* 8.12, 1807–1827. DOI: 10.1080/01431168708954818. URL: http://www.tandfonline.com/doi/abs/10.1080/01431168708954818.

Mokvist, F., F. Mamedov, and S. Styring (2014). "Defining the Far-red limit of photosystem I: The primary charge separation is functional to 840 nm". *Journal of Biological Chemistry* 289.35, 24630–24639. DOI: 10.1074/jbc.M114.555649. URL: http://www.jbc.org/lookup/doi/10.1074/jbc.M114.555649.

Mondino, E. B. (2018). "Advances in Service and Industrial Robotics". 49. DOI: 10.1007/978-3-319-61276-8. URL: http://link.springer.com/10.1007/978-3-319-61276-8.

Morales, A., R. Guerra, P. Horstrand, M. Diaz, A. Jimenez, J. Melian, S. Lopez, and J. F. Lopez (2020). "A Multispectral Camera Development: From the Prototype Assembly until Its Use in a UAV System". *Sensors* 20.21, 6129. DOI: 10.3390/s20216129. URL: https://www.mdpi.com/1424-8220/20/21/6129.

Morimoto, S., V. Scholar, and W. F. McClure (1999). "More on Derivatives: Resolving Overlapping Absorbance Bands". *NIR news* 10.4, 10–12. DOI: 10.1255/nirn.529. URL: http://journals.sagepub.com/doi/10.1255/nirn.529.

Muñoz, J. D., A. O. Finley, R. Gehl, and S. Kravchenko Sasha (2010). "Nonlinear hierarchical models for predicting cover crop biomass using Normalized Difference Vegetation Index". *Remote Sensing of Environment* 114.12, 2833–2840. DOI: 10.1016/j.rse.2010.06.011. URL: http://dx.doi.org/10.1016/j.rse.2010.06.011.

Mutanga, O. (2004). "Hyperspectral remote sensing of tropical grass quality and quantity". PhD thesis. Twente, 195.

Mutanga, O. and A. K. Skidmore (2004a). "Narrow band vegetation indices overcome the saturation problem in biomass estimation". *International Journal of Remote Sensing* 25.19, 3999–4014. DOI: 10.1080/01431160310001654923. URL: https://www.tandfonline.com/doi/full/10.1080/01431160310001654923.

Mutanga, O. and A. K. Skidmore (2004b). "Hyperspectral band depth analysis for a better estimation of grass biomass (Cenchrus ciliaris) measured under controlled laboratory conditions". *International Journal of Applied Earth Observation and Geoinformation* 5.2, 87–96. DOI: 10.1016/j.jag.2004.01.001.

– (2007). "Red edge shift and biochemical content in grass canopies". *ISPRS Journal of Photogrammetry and Remote Sensing* 62.1, 34–42. DOI: 10.1016/j.isprsjprs.2007.02.001. URL: https://linkinghub.elsevier.com/retrieve/pii/S0924271607000068.

Nakagami, K. (2016). "Effects of sites and years on the coefficients of rising plate meter calibration under varying coefficient models". *Grassland Science* 62.2, 128–132. DOI: 10.1111/grs.12117. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/grs.12117%20http://doi.wiley.com/10.1111/grs.12117.

Norris, K. H., R. F. Barnes, J. E. Moore, and J. S. Shenk (1976). "Predicting Forage Quality by Infrared Reflectance Spectroscopy". *Journal of Animal Science* 43.4, 889–897. DOI: 10.2527/jas1976.434889x. URL: https://academic.oup.com/jas/article/43/4/889-897/4697632.

Oenema, J., M. van Ittersum, and H. van Keulen (2012). "Improving nitrogen management on grassland on commercial pilot dairy farms in the Netherlands". *Agriculture, Ecosystems and Environment* 162.2012, 116–126. DOI: 10.1016/j.agee.2012.08.012. URL: http://dx.doi.org/10.1016/j.agee.2012.08.012.

Oliveira, R. A., R. Näsi, O. Niemeläinen, L. Nyholm, K. Alhonoja, J. Kaivosoja, L. Jauhiainen, N. Viljanen, S. Nezami, L. Markelin, T. Hakala, and E. Honkavaara (2020). "Machine learning estimators for the quantity and quality of grass swards used for silage production using drone-based imaging spectrometry and photogrammetry". *Remote Sensing of Environment* 246.December 2019, 111830. DOI: 10.1016/j.rse.2020.111830. URL: https://doi.org/10.1016/j.rse.2020.111830.

Opitz, D. and R. Maclin (1999). "Popular Ensemble Methods: An Empirical Study". *Journal of Artificial Intelligence Research* 11, 169–198. DOI: 10.1613/jair.614. URL: https://jair.org/index.php/jair/article/view/10239.

Paredes, J. A., J. Gonzalez, C. Saito, and A. Flores (2017). "Multispectral imaging system with UAV integration capabilities for crop analysis". *IEEE 1st International Symposium on Geoscience and Remote Sensing, GRSS-CHILE 2017*. DOI: 10.1109/GRSS-CHILE.2017.7996009.

Pearson, R. L., L. D. Miller, and C. J. Tucker (1976). "Hand held spectral radiometer to estimate gramineous biomass". *Applied Optics* 15.2, 416–419. DOI: 10.1364/AO.15.000416.

Peñuelas, J. and I. Filella (1995). "Reflectance assessment of mite effects on apple trees". *International Journal of Remote Sensing* 16.14, 2727–2733. DOI: 10.1080/01431169508954588.

Perez-Riverol, Y., M. Kuhn, J. A. Vizcaíno, M. P. Hitz, and E. Audain (2017). "Accurate and fast feature selection workflow for high-dimensional omics data". *PLoS ONE* 12.12, 1–14. DOI: 10.1371/journal.pone.0189875.

Perry, C. R. and L. F. Lautenschlager (1984). "Functional equivalence of spectral vegetation indices". *Remote Sensing of Environment* 14.1-3, 169–182. DOI: 10.1016/0034-4257(84)90013-0.

Pircher, M., J. Geipel, K. Kusnierek, and A. Korsaeth (2017). "Development of a hybrid UAV sensor platform suitable for farm-scale applications in precision agriculture". *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-2/W6.September, 297–302. DOI: 10.5194/isprs-archives-XLII-2-W6-297-2017. URL: https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W6/297/2017/.

Poncet, A. M., T. Knappenberger, C. Brodbeck, M. Fogle, J. N. Shaw, and B. V. Ortiz (2019). "Multispectral UAS Data Accuracy for Different Radiometric Calibration Methods". *Remote Sensing* 11.16, 1917. DOI: 10.3390/rs11161917. URL: https://www.mdpi.com/2072-4292/11/16/1917.

Prewer, W. E., A. V. Bysterveldt, A. v. B. .-. o. t. N. Z. G. ..., undefined 2004, and A. V. Bysterveldt (2004). "The Pasture Quality Poster-a learning tool for farmers". In: *Proceedings of the New Zealand Grassland Association*, 183–186. URL: https://www.grassland.org.nz/publications/nzgrassland%7B%5C_%7Dpublication%7B%5C_%7D435.pdf.

Price, J. C. (1994). "Band selection procedure for multispectral scanners". *Applied Optics* 33.15, 3281. DOI: 10.1364/ao.33.003281.

Pricope, N. G., K. L. Mapes, K. D. Woodward, S. F. Olsen, and J. B. Baxley (2019). "Multi-Sensor Assessment of the Effects of Varying Processing Parameters on UAS Product Accuracy and Quality". *Drones* 3.3, 63. DOI: 10.3390/drones3030063. URL: https://www.mdpi.com/2504-446X/3/3/63.

Pullanagari, R. R., G. Kereszturi, I. J. Yule, and M. E. Irwin (2015). "Determination of pasture quality using airborne hyperspectral imaging". *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVII* 9637.April 2016, 963706. DOI: 10.1117/12.2193844.

Pullanagari, R. R., G. Kereszturi, and I. J. Yule (2016). "Mapping of macro and micro nutrients of mixed pastures using airborne AisaFENIX hyperspectral imagery". *ISPRS Journal of Photogrammetry and Remote Sensing* 117, 1–10. DOI: 10.1016/j.isprsjprs.2016.03.010. URL: http://www.sciencedirect.com/science/article/pii/S0924271616000782.

Pullanagari, R. R., I. J. Yule, M. P. Tuohy, M. J. Hedley, R. A. Dynes, and W. M. King (2012). "In-field hyperspectral proximal sensing for estimating quality parameters of mixed pasture". *Precision Agriculture* 13.3, 351–369. DOI: 10.1007/s11119-011-9251-4.

Pullanagari, R. R. (2011). "Proximal sensing techniques to monitor pasture quality and quantity on dairy farms". en. PhD thesis. URL: http://hdl.handle.net/10179/3695.

Quinlan, J. R. (1992). "Learning with continuous classes". In: *Australian Joint Conference on Artificial Intelligence*. Vol. 92, 343–348.

– (1993). "Combining Instance-Based and Model-Based Learning". In: *Machine Learning Proceedings 1993*. Vol. 93. Elsevier, 236–243. DOI: 10.1016/B978-1-55860-307-3.50037-X. URL: https://linkinghub.elsevier.com/retrieve/pii/B978155860307350037X.

R Core Team and R Development Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. URL: http://www.r-project.org%20https://www.r-project.org/.

Ramoelo, A., A. K. Skidmore, M. Schlerf, R. Mathieu, and I. M. Heitkönig (2011). "Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations". *ISPRS Journal of Photogrammetry and Remote Sensing* 66.4, 408–417. DOI: 10.1016/j.isprsjprs.2011.01.008. URL: http://dx.doi.org/10.1016/j.isprsjprs.2011.01.008%20https://linkinghub.elsevier.com/retrieve/pii/S0924271611000232.

Rasmussen, M. A. and R. Bro (2012). "A tutorial on the Lasso approach to sparse modeling". *Chemometrics and Intelligent Laboratory Systems* 119.November 2019, 21–31. DOI: 10.1016/j.chemolab.2012.10.003. URL: http://dx.doi.org/10.1016/j.chemolab.2012.10.003.

Rawnsley, R. P., A. D. Langworthy, K. G. Pembleton, L. R. Turner, R. Corkrey, and D. J. Donaghy (2014). "Quantifying the interactions between grazing interval, grazing intensity, and nitrogen on the yield and growth rate of dryland and irrigated perennial ryegrass". *Crop and Pasture Science* 65.8, 735–746. DOI: 10.1071/CP13453.

Regan, Á., J. Douglas, J. Maher, and T. O'Dwyer (2020). "Exploring farmers' decisions to engage in grass measurement on dairy farms in Ireland". *Journal of Agricultural Education and Extension* 0.0, 1–26. DOI: 10.1080/1389224X.2020.1858892. URL: https://doi.org/10.1080/1389224X.2020.1858892.

Roberts, C. A., J. Stuth, and P. Flinn (2004). "Analysis of Forages and Feedstuffs". In: *Near-infrared spectroscopy in agriculture*. Vol. 44. Hoboken,NJ,USA: Wiley Online Library, 229–267. DOI: 10.2134/agronmonogr44.c10. URL: http://doi.wiley.com/10.2134/agronmonogr44.c10.

Roberts, D. R., V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, D. I. Warton, B. A. Wintle, F. Hartig, and C. F. Dormann (2017). "Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure". *Ecography* 40.8, 913–929. DOI: 10.1111/ecog.02881. URL: http://doi.wiley.com/10.1111/ecog.02881.

Rohantgi, A. (2020). *WebPlotDigitizer*. Pacifica, Callehifornia. URL: https://automeris.io/WebPlotDigitizer.

Roscher, R., B. Bohn, M. F. Duarte, and J. Garcke (2020). "Explain it to me-facing remote sensing challenges in the bio-and geosciences with explainable machine learning". *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 5.3, 817–824. DOI: 10.5194/isprs-Annals-V-3-2020-817-2020.

Rouse, J. W., R. H. Hass, J. Schell, and D. Deering (1973). "Monitoring vegetation systems in the great plains with ERTS". *Third Earth Resources Technology Satellite (ERTS) symposium* 1, 309–317. URL: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022614.pdf.

Sanches, I., M. Tuohy, M. Hedley, and A. Mackay (2013). "Seasonal prediction of in situ pasture macronutrients in New Zealand pastoral systems using hyperspectral data". *International Journal of Remote Sensing* 34.1, 276–302. DOI: 10.1080/01431161.2012.713528. URL: https://www.tandfonline.com/doi/full/10.1080/01431161.2012.713528.

Sanches, I. D. A. (2009). "Hyperspectral proximal sensing of the botanical composition and nutrient content of New Zealand pastures", 219. URL: http://hdl.handle.net/10179/1194.

Sandmeier, S. and K. Itten (1999). "A field goniometer system (FIGOS) for acquisition of hyperspectral BRDF data". *IEEE Transactions on Geoscience and Remote Sensing* 37.2, 978–986. DOI: 10.1109/36.752216. URL: http://ieeexplore.ieee.org/document/752216/.

Schapire, R. E. (1990). "The Strength of Weak Learnability". 227, 197–227.

Schellberg, J., M. J. Hill, R. Gerhards, M. Rothmund, and M. Braun (2008). "Precision agriculture on grassland: Applications, perspectives and constraints". *European Journal of Agronomy* 29.2-3, 59–71. DOI: 10.1016/j.eja.2008.05.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S1161030108000646.

Schiller, S. J. and J. Silny (2010). "The Specular Array Radiometric Calibration (SPARC) method: a new approach for absolute vicarious calibration in the solar reflective spectrum". *Remote Sensing System Engineering III* 7813.August 2010, 78130E. DOI: 10.1117/12.864071.

Scrucca, L. (2013). "GA: A package for genetic algorithms in R". *Journal of Statistical Software* 53.4, 1–37. DOI: 10.18637/jss.v053.i04.

Sellers, P. J. (1985). "Canopy reflectance, photosynthesis and transpiration". *International Journal of Remote Sensing* 6.8, 1335–1372. DOI: 10.1080/01431168508948283.

Shalloo, L., M. O. Donovan, L. Leso, J. Werner, E. Ruelle, A. Geoghegan, L. Delaby, and N. O. Leary (2018). "Review: Grass-based dairy systems, data and precision

technologies". *Animal* 12.s2, S262–S271. DOI: 10.1017/S175173111800246X. URL: http://dx.doi.org/10.1017/S175173111800246X.

Shorten, P. R., S. R. Leath, J. Schmidt, and K. Ghamkhar (2019). "Predicting the quality of ryegrass using hyperspectral imaging". *Plant Methods* 15.1, 63. DOI: 10.1186/s13007-019-0448-2. URL: https://plantmethods.biomedcentral.com/articles/10.1186/s13007-019-0448-2.

Silleos, N. G., T. K. Alexandridis, I. Z. Gitas, and K. Perakis (2006). "Vegetation indices: Advances made in biomass estimation and vegetation monitoring in the last 30 years". *Geocarto International* 21.4, 21–28. DOI: 10.1080/10106040608542399.

Starks, P. J. and M. A. Brown (2010). "Prediction of Forage Quality from Remotely Sensed Data: Comparison of Cultivar-Specific and Cultivar-Independent Equations Using Three Methods of Calibration". *Crop Science* 50.5, 2159–2170. DOI: 10.2135/cropsci2009.08.0455. URL: http://doi.wiley.com/10.2135/cropsci2009.08.0455.

Starks, P. J., M. A. Brown, K. E. Turner, and B. C. Venuto (2016). "Canopy Visible and Near-infrared Reflectance Data to Estimate Alfalfa Nutritive Attributes Before Harvest". *Crop Science* 56.1, 484–496. DOI: 10.2135/cropsci2015.03.0162. URL: http://doi.wiley.com/10.2135/cropsci2015.03.0162.

Starks, P. J., D. Zhao, W. A. Phillips, and S. W. Coleman (2006). "Development of Canopy Reflectance Algorithms for Real-Time Prediction of Bermudagrass Pasture Biomass and Nutritive Values". *Crop Science* 46.2, 927–934. DOI: 10.2135/cropsci2005.0258. URL: https://onlinelibrary.wiley.com/doi/abs/10.2135/cropsci2005.0258.

Stone, M. (1974). "Cross-Validatory Choice and Assessment of Statistical Predictions". *Journal of the Royal Statistical Society. Series B (Methodological)* 36.2, 111–147. arXiv: /www.jstor.org/stable/2984809 [https:]. URL: http://www.jstor.org/stable/2984809.

Suomalainen, J., T. Hakala, R. A. de Oliveira, L. Markelin, N. Viljanen, R. Näsi, and E. Honkavaara (2018). "A novel tilt correction technique for irradiance sensors and spectrometers on-board unmanned aerial vehicles". *Remote Sensing* 10.12, 1–18. DOI: 10.3390/rs10122068.

Thenkabail, P. S., R. B. Smith, and E. De Pauw (2002). "Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization". *Photogrammetric Engineering and Remote Sensing* 68.6, 607–621.

Thomson, A. L., S. B. Karunaratne, A. Copland, D. Stayches, E. M. McNabb, and J. Jacobs (2020). "Use of traditional, modern, and hybrid modelling approaches for in situ prediction of dry matter yield and nutritive characteristics of pasture using hyperspectral datasets". *Animal Feed Science and Technology* 269.September, 114670. DOI: 10.1016/j.anifeedsci.2020.114670. URL: https://linkinghub.elsevier.com/retrieve/pii/S0377840120305745.

Thomson, N. A., D. A. Mccallum, S. Howse, C. W. Holmes, P. N. P. Matthews, and C. Matthew (1997). "Estimation of dairy pastures - the need for standardisation". In: *Proceedings of the New Zealand Grassland Association*. Vol. 59. 59, 221–225.

Thulin, S. (2008). "Hyperspectral remote sensing of temperate pasture quality". PhD thesis. RMIT University. URL: http://researchbank.rmit.edu.au/view/rmit:7867.

Thulin, S., M. J. Hill, A. Held, S. Jones, and P. Woodgate (2014). "Predicting Levels of Crude Protein, Digestibility, Lignin and Cellulose in Temperate Pastures Using Hyperspectral Image Data". *American Journal of Plant Sciences* 05.07, 997–1019. DOI: 10.4236/ajps.2014.57113. URL: http://www.scirp.org/journal/doi.aspx?DOI=10.4236/ajps.2014.57113.

Togeiro de Alckmin, G., L. Kooistra, R. Rawnsley, S. de Bruin, and A. Lucieer (2020a). "Retrieval of Hyperspectral Information from Multispectral Data for Perennial Ryegrass Biomass Estimation". *Sensors* 20.24, 7192. DOI: 10.3390/s20247192. URL: https://www.mdpi.com/1424-8220/20/24/7192.

Togeiro de Alckmin, G., L. Kooistra, R. Rawnsley, and A. Lucieer (2021). "Comparing methods to estimate perennial ryegrass biomass: canopy height and spectral vegetation indices". *Precision Agriculture* 22.1, 205–225. DOI: 10.1007/s11119-020-09737-z. URL: http://link.springer.com/10.1007/s11119-020-09737-z.

Togeiro de Alckmin, G., A. Lucieer, G. Roerink, R. Rawnsley, I. Hoving, and L. Kooistra (2020b). "Retrieval of Crude Protein in Perennial Ryegrass Using Spectral Data at the Canopy Level". *Remote Sensing* 12.18, 2958. DOI: 10.3390/rs12182958. URL: https://www.mdpi.com/2072-4292/12/18/2958.

Tucker, C. J., L. D. Miller, and R. L. Pearson (1973). "Measurement of the combines effect of green biomass, chlorophyll and leaf water on canopy spectroreflectance of the shortgrass prairie". English. In: *Proceedings of the second annual remote sensing of Earth Resources Conference*. Fort Collins: Space Institute University of Tennessee. URL: https://books.google.nl/books?id=tu8sAQAAMAAJ.

– (1975). "Shortgrass prairie spectral measurements. [for terrain analysis and photomapping]". *Photogrammetric Engineering and Remote Sensing* 41, 1157–1162.

Tucker, C. J. (1977a). "Asymptotic nature of grass canopy spectral reflectance". *Applied Optics* 16.5, 1151. DOI: 10.1364/AO.16.001151. URL: https://www.osapublishing.org/abstract.cfm?URI=ao-16-5-1151.

– (1977b). "Spectral estimation of grass canopy variables". *Remote Sensing of Environment* 6.1, 11–26. DOI: 10.1016/0034-4257(77)90016-5. URL: https://linkinghub.elsevier.com/retrieve/pii/0034425777900165.

– (1979). "Red and photographic infrared linear combinations for monitoring vegetation". *Remote Sensing of Environment* 8.2, 127–150. DOI: 10.1016/0034-4257(79)90013-0. arXiv: arXiv:1011.1669v3.

Turner, D., A. Lucieer, and C. Watson (2012). "An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SfM) Point Clouds". *Remote Sensing* 4.5, 1392–1410. DOI: 10.3390/rs4051392. URL: http://www.mdpi.com/2072-4292/4/5/1392.

Turner, L. R., D. J. Donaghy, P. A. Lane, and R. P. Rawnsley (2006). "Effect of defoliation management, based on leaf stage, on perennial ryegrass (Lolium perenne L.), prairie grass (Bromus willdenowii Kunth.) and cocksfoot (Dactylis glomerata L.) under dryland conditions. 2. Nutritive value". *Grass and Forage Science* 61.2, 175–181. DOI: 10.1111/j.1365-2494.2006.00524.x. URL: http://doi.wiley.com/10.1111/j.1365-2494.2006.00524.x.

Valavi, R., J. Elith, J. J. Lahoz-Monfort, and G. Guillera-Arroita (2019). "blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models". *Methods in Ecology and Evolution* 10.2, 225–232. DOI: 10.1111/2041-210X.13107.

Valk, H., I. E. Leusink-Kappers, and A. M. Van Vuuren (2000). "Effect of reducing nitrogen fertilizer on grassland on grass intake, digestibility and milk production of dairy cows". *Livestock Production Science* 63.1, 27–38. DOI: 10.1016/S0301-6226(99)00118-9.

Verrelst, J., G. Camps-Valls, J. Muñoz-Marí, J. P. Rivera, F. Veroustraete, J. G. P. W. Clevers, and J. Moreno (2015). "Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - A review". *ISPRS Journal of Photogrammetry and Remote Sensing* 108, 273–290. DOI: 10.1016/j.isprsjprs.2015.05.005. URL: http://dx.doi.org/10.1016/j.isprsjprs.2015.05.005.

Verstraete, M. M., B. Pinty, and R. B. Myneni (1996). "Potential and limitations of information extraction on the terrestrial biosphere from satellite remote sensing". *Remote Sensing of Environment* 58.2, 201–214. DOI: 10.1016/S0034-4257(96)00069-7.

Vescovo, L., G. Wohlfahrt, M. Balzarolo, S. Pilloni, M. Sottocornola, M. Rodeghiero, and D. Gianelle (2012). "New spectral vegetation indices based on the near-infrared shoulder wavelengths for remote detection of grassland phytomass". *International Journal of Remote Sensing* 33.7, 2178–2195. DOI: 10.1080/01431161.2011.607195. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3859895%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract%20https://www.tandfonline.com/doi/full/10.1080/01431161.2011.607195.

Von Bueren, S. K., A. Burkart, A. Hueni, U. Rascher, M. P. Tuohy, and I. J. Yule (2015). "Deploying four optical UAV-based sensors over grassland: Challenges and limitations". *Biogeosciences* 12.1, 163–175. DOI: 10.5194/bg-12-163-2015.

Wachendorf, M., M. Buchter, H. Trott, and F. Taube (2004). "Performance and environmental effects of forage production on sandy soils. II. Impact of defoliation system and nitrogen input on nitrate leaching losses". *Grass and Forage Science* 59.1,

56–68. DOI: 10.1111/j.1365-2494.2004.00401.x. URL: http://doi.wiley.com/10.1111/j.1365-2494.2004.00401.x.

Wang, J., P. Badenhorst, A. Phelan, L. Pembleton, F. Shi, N. Cogan, G. Spangenberg, and K. Smith (2019). "Using Sensors and Unmanned Aircraft Systems for High-Throughput Phenotyping of Biomass in Perennial Ryegrass Breeding Trials". *Frontiers in Plant Science* 10.October, 1–9. DOI: 10.3389/fpls.2019.01381.

Ward, J. H. (1963). "Hierarchical Grouping to Optimize an Objective Function". *Journal of the American Statistical Association* 58.301, 236–244. DOI: 10.1080/01621459.1963.10500845. URL: http://www.tandfonline.com/doi/abs/10.1080/01621459.1963.10500845.

Weiss, M., F. Jacob, and G. Duveiller (2020). "Remote sensing for agricultural applications: A meta-review". *Remote Sensing of Environment* 236.December 2018, 111402. DOI: 10.1016/j.rse.2019.111402. URL: https://doi.org/10.1016/j.rse.2019.111402.

Westad, F. and F. Marini (2015). "Validation of chemometric models - A tutorial". *Analytica Chimica Acta* 893.April 2019, 14–24. DOI: 10.1016/j.aca.2015.06.056. URL: http://dx.doi.org/10.1016/j.aca.2015.06.056.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani (2019). "Welcome to the {tidyverse}". *Journal of Open Source Software* 4.43, 1686. DOI: 10.21105/joss.01686.

Wiersma, D. J. and D. A. Landgrebe (1980). "Analytical design of multispectral sensors". *IEEE Transactions on Geoscience and Remote Sensing* GE-18.2, 180–189. DOI: 10.1109/TGRS.1980.350271.

Wijesingha, J., T. Astor, D. Schulze-Brüninghoff, M. Wengert, and M. Wachendorf (2020). "Predicting Forage Quality of Grasslands Using UAV-Borne Imaging Spectroscopy". *Remote Sensing* 12.1, 126. DOI: 10.3390/rs12010126. URL: https://www.mdpi.com/2072-4292/12/1/126.

Wilkinson, J. M., M. R. Lee, M. J. Rivero, and A. T. Chamberlain (2020). "Some challenges and opportunities for grazing dairy cows on temperate pastures". *Grass and Forage Science* 75.1, 1–17. DOI: 10.1111/gfs.12458.

Wold, S., M. Sjöström, and L. Eriksson (2001). "PLS-regression: A basic tool of chemometrics". *Chemometrics and Intelligent Laboratory Systems* 58.2, 109–130. DOI: 10.1016/S0169-7439(01)00155-1.

Xue, J. and B. Su (2017). "Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications". *Journal of Sensors* 2017, 1–17. DOI: 10.1155/2017/1353691. URL: https://www.hindawi.com/journals/js/2017/1353691/.

Zalles, V., M. C. Hansen, P. V. Potapov, S. V. Stehman, A. Tyukavina, A. Pickens, X. P. Song, B. Adusei, C. Okpa, R. Aguilar, N. John, and S. Chavez (2019). "Near doubling of Brazil's intensive row crop area since 2000". *Proceedings of the National Academy of Sciences of the United States of America* 116.2, 428–435. DOI: 10.1073/pnas.1810301115.

Zarco-Tejada, P. J., J. C. Pushnik, S. Dobrowski, and S. L. Ustin (2003). "Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects". *Remote Sensing of Environment* 84.2, 283–294. DOI: 10.1016/S0034-4257(02)00113-X.

Zebarth, B. J., M. Younie, J. W. Paul, and S. Bittman (2002). "Evaluation of leaf chlorophyll index for making fertilizer nitrogen recommendations for silage corn in a high fertility environment". *Communications in Soil Science and Plant Analysis* 33.5-6, 665–684. DOI: 10.1081/CSS-120003058. URL: https://www.tandfonline.com/doi/full/10.1081/CSS-120003058.

Zeng, C., D. J. King, M. Richardson, and B. Shan (2017). "Fusion of multispectral imagery and spectrometer data in UAV remote sensing". *Remote Sensing* 9.7. DOI: 10.3390/rs9070696.

Zhao, J., X. Zhang, C. Gao, X. Qiu, Y. Tian, Y. Zhu, and W. Cao (2019). "Rapid mosaicking of unmanned aerial vehicle (UAV) images for crop growth monitoring using the SIFT algorithm". *Remote Sensing* 11.10. DOI: 10.3390/rs11101226.

# Summary

Pastures are the cornerstone of grazing-based livestock production systems, allowing for sustainable use of marginal agricultural lands, while transforming a non-digestible resource into nutritious products fit for human consumption. Despite the extensive literature about the constituent factors of pasture production and management, the majority of pasture-based grazing systems are not optimally managed due to the costly and time-consuming nature of current methods for measurement and monitoring. Although challenging, outdoor spectral observations have the potential to provide real-time information about pasture quantity and quality, serving as an ideal sensing technique for autonomous platforms, consequently, alleviating the monitoring bottleneck and supporting intensive pasture production.

The overarching aim of this thesis is to determine to which extent spectral data , in outdoor environments, can accurately estimate key biophysical and biochemical components of perennial ryegrass (Lolium perenne). In addition, this study aims to validate the transferability of spectral models from handheld spectral measurements to remotely piloted aerial systems (RPAS), while critically assessing optimal modelling approaches and minimal sensor requirements. In doing so, this thesis validates the superiority of spectral data over commonly practiced canopy-height models (CHM) for biomass estimation and suggests that suitable low-cost imaging sensing systems are within commercial reach.

After the Introduction, chapter two critically addresses the use of vegetation indices and canopy height for biomass estimation, while quantifying accuracy improvements based on different regression algorithms. As a proxy for status-quo techniques, a comparison between CHM and the normalized vegetation index (NDVI) is performed. In addition, to further explore the potential of vegetation indices , a brute-force procedure was employed to generate 11,026 normalized ratio indices (NRI) while selecting the best NRI band combination. In parallel, a pool 97 literature based vegetation indices, was filtered and underwent a feature selection procedure to determine an optimal small subset of indices. Results suggest that: (i) an optimized vegetation index (i.e., best NRI) and CHM are equivalent; (ii) a small number of vegetation indices is sufficient to reach achievable accuracy when employing top-of-canopy reflectance alone; and (iii) accuracies and precision can be improved solely through more elaborate modelling techniques, such as non-parametric

methods and model stacking. In chapter three, a genetic algorithm was employed in a two-objective search procedure: to minimize the number of spectral bands while simultaneously maximizing model accuracy for crude protein estimation. This protocol was employed over different spectral ranges, namely VIS-NIR, SWIR and the Full- Spectrum range, while comparing achievable accuracies of two different metrics: crude protein as dry matter fraction (% CP) or in a weight-per-area basis (kg CP/ha). Results suggest that, in outdoor environments, the best approach to estimate crude protein relies on its expression as weight-per-area basis and that the VIS-NIR alone can provide best accuracies in both known and unseen locations.

Chapter four presents a new approach for retrieval of a continuous spectral signature (550-790 nm) from discrete multispectral measurements (i.e., four bands, as per a commercially available multispectral camera) based on a piecewise function described by two parametric sub-functions. The retrieval of spectral signatures allowed for the generation of continuum-removed features and associated vegetation indices for prediction of biomass, previously reported in the literature as optimal indices for biomass estimation . These synthetic vegetation indices were compared against vegetation indices derived from the original band values. No significant improvement in performance was found, suggesting that underlying biological broadband absorption features (e.g. pigments, leaf-area and cell structure) were well described by both reflectance-based and continuum-removed features. Consequently, it is suggested that achievable accuracy is largely driven by an appropriate model fit between any spectral metrics of these broadband absorption features.

Chapter five employed RPAS multispectral imagery, handheld spectral-data and five distinct decision-rule regression techniques to validate the approach of biomass assessment employing a small subset of indices, while critically addressing the challenges in radiometric calibration, model interpretability, model deployment in an operational scenario, and model-performance through different validation strategies. The five regression algorithms build upon the concept of regression-trees, using techniques of bootstrapping aggregation (i.e., bagging) and boosting, consequently increasing model complexity (e.g., number of trees, depth of trees) while decreasing overall interpretability.

The RPAS multispectral was compared against handheld top-of-the canopy spectral measurements and significant inconsistencies were found between reflectance values of both sensors. Consequently, this chapter indicates the absence of well-defined and robust protocols for spectral data collection of commercial multispectral cameras When calibrated through a thorough pipeline, multispectral data was able to provide better results, although with average marginal superiority (i.e. 60kg) than handheld data.

When employing a repeated k-fold cross-validation for model associated with the multispectral imagery, two distinct algorithms (i.e., Cubist and Random-Forest) presented equivalent performances , presenting an equivalent error-metric of nearly 400 kg DM/ha. However, the three remaining (i.e., bagged, boosted trees and CART) algorithms had

an average performance of 450 kg DM/ha. Both CART and Cubist are considered more interpretable, operating under a single regression tree structure, thus, presenting a faster prediction speed, smaller size and the ability to scrutinize models. However, a temporal validation strategy showed a low reliability of spectral models when validated outside its boundary conditions, with associated errors above 800 kg DM/ha, rendering predictions not useful and showcasing the shortcomings of performance claims of short-duration studies.

In its Outlook, special emphasis is directed to maximum achievable accuracies through the use of spectral data in outdoor environments, under the presence of confounding and masking effects derived from canopy geometry and illumination conditions while stressing the need for rigorous protocols and quality assurance mechanisms for spectral data collection. Finally, this thesis identifies the main bottlenecks for the advancement of spectral imaging techniques in a farm-operational scenario, indicating possible advancement through the use of minimal sensing equipment, automated data-collection, faster and interpretable modelling techniques.

# Acknowledgements

I would like to express my gratitude towards my supervisors: Prof. Arko Lucieer, Prof. Lammert Kooistra, and Dr. Richard Rawnsley. Also, I am grateful for the contributions provided by my co-authors: Prof. Sytze de Bruin, Idse Hoving, and Gerbert Roerink.

I sincerely thank the contributions made by ASD PanAnalytical and Eurofins, providing equipment and services through in-kind donations. This program would not have been possible without the support of Australian dairy farmers and their levy. I hope that the content presented in this thesis has merited such investment.

I would like to extend my gratitude to my colleagues and friends in Tasmania, the Netherlands, and Brazil.

Finally, to my parents: without your support, this would not have been possible.

Gustavo
Wageningen
January 10$^{th}$, 2021

# About the author

Gustavo Togeiro de Alckmin was born on 17 May 1987 in São Paulo, Brazil. In 2005, Gustavo studied Agronomic Engineering at the University São Paulo (Campus "Luiz de Queiroz" - Piracicaba) from 2005 to 2010. In 2010 he started his professional careeer as a Sugar Trader in a British trading company, providing pricing and hedging advisory.

In 2011, he completed a *Lato Sensu* specialization in Future Markets and Derivatives at the Brazilian Mercantile Future and Exchange. From 2012 to 2014, he attendend a double-degree Master of Science program at the Universities of Montpellier SupAgro and Universidad Politecnica de Madrid. His Master Thesis dissertation was about the use of unmanned aerial vehicle and remote sensing techniques for pasture management. During the years of 2014 and 2015, he worked in his own startup company in the same topic.

From 2016 onwards, he joined the doctoral program at the University of Tasmania and Wageningen University to further explore the topic of remote sensing for pasture management.

# PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)


**Review of Literature (6 ECTS)**

- Remote Sensing Techniques for Pasture Measurement and Monitoring – A general overview and emphasis on Remotely Piloted Aircraft Systems.

**Post-graduate courses (7.7 ECTS)**

- Dairy Production and Management (online); ); PennState Uni (2016)
- Chemometrics; VLAG (2018)
- Mixed Linear Models; PE&RC (2018)
- Machine Learning for Spatial Data; PE&RC (2018)
- Generalized Linear Models; PE&RC (2018)
- Experimental Design; PE&RC (2018)
- Meta–Analaysis; PE&RC (2018)
- Summer School Open GeoHub; Open GeoHub (2019)

**Deficiency, refresh, brush-up courses (1.8 ECTS)**

- Chemooc Basics: the basis of chemometrics; Agreenium (2017)
- Chemooc advanced: advanced chemometrics, data validations; Agreenium (2017)

**Competence strengthening / skills courses (2.1 ECTS)**

- Introduction to Higher Degree by Research 1$^{st}$ part; UTAS (2016)

- Communicating Research; PE&RC (2016)
- High-Impact Writing in Science; PE&RC (2019)
- Survival guide to peer-review; PE&RC (2019)

**Scientific Integrity/Ethics in science activities (0.3 ECTS)**

- Introduction to Higher Degree by Research 2nd part; UTAS

**PE&RC Annual meetings, seminars and the PE&RC weekend (1.2 ECTS)**

- PE&RC First years weekend (2018)
- PE&RC Last years weekend (2019)

**National scientific meetings / local seminars / discussion groups (4.7 ECTS)**

- Writing Workshop; UTAS SLAF (2016)
- Software Carpentry Hobart (2016)
- Writing Retreat Hobart (2017)
- Tasmanian Farmer Discussion Group (2017–2018)
- Tasmanian Dairy Conference (2017–2018)
- R Discussion group (2018–19)
- Precision farming – using new technologies to optimise grassland systems (2019)

**International symposia, workshops and conferences (5.8 ECTS)**

- Australian Dairy Conference; Australia (2017)
- Unmanned Aerial Vehicles for Remote Sensing Conference; Australia (2017)
- European Conference for Precision Agriculture; France (2019)
- GeoSpatial Week (Hyperspectral meets Machine Learning); The Netherlands (2019)

**Societally relevant exposure (0.9 ECTS)**

- Farming Ahead - magazine (2017)
- AgFest Launceston Australia – tech exhibition (2017)
- AgInstitute Devonport. Australia – guest lecturer (2017)

**Lecturing / supervision of practical's / tutorials (3.3 ECTS)**

- Introduction to Remote Sensing
- Advanced Earth Observation

**BSc / MSc thesis supervision (3 ECTS)**

- Remote sensing Hollywood: Computer Generated Forests

This research was supported by Dairy Australia, through the Dairy on PAR action.

Cover design by Gustavo Togeiro de Alckmin