

Perennial ryegrass biomass retrieval through multispectral UAV data

Gustavo Togeirode Alckmin^{d,*}, Arko Lucieer^b, Richard Rawnsley^{b,c}, Lammert Kooistra^a

^a University of Tasmania - Discipline of Geography and Spatial Sciences, School of Geography, Planning, and Spatial Sciences, Hobart, TAS 7005, Australia

^b Wageningen University - Laboratory of Geo-Information Science and Remote Sensing, Droevendaalsesteeg 3, 6708 PB Wageningen, the Netherlands

^c Tasmanian Institute of Agriculture - Centre for Dairy, Grains and Grazing, 16-20 Mooreville Road, Burnie, TAS 7320, Australia

^d University of Missouri - College of Agriculture, Food and Natural Resources, Columbia, MO 67201, United States

ARTICLE INFO

Keywords:

Perennial ryegrass
Machine learning
Vegetation indices
Biomass
UAV
Radiometric calibration

ABSTRACT

Frequent biomass measurement is a key activity for optimal perennial ryegrass (*Lolium perenne*) management in intensive forage-based dairy operations. Due to the necessary high-frequency (i.e., weekly or monthly) pasture monitoring and continuous trend of larger dairy farms, such activity is perceived as an operational bottleneck. Consequently, substantial effort is directed to the development of accurate and automated technological solutions for biomass assessment. The 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. Nevertheless, the accuracy of radiometric calibration and generalization ability of these models is seldom evaluated. Also, 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. 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 (i.e., 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.

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. The recent widespread adoption of *Unmanned Aerial Vehicles (UAVs)* 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, can now conveniently capitalize on *Remote Sensing (RS)* techniques and semi-autonomous aerial platforms for biomass assessment.

In the past decades, substantial progress has been achieved in determining key biophysical attributes of pastures (e.g., biomass) through the analysis of spectral data, particularly through the use of field spectrometers (Mutanga and Skidmore, 2004). Although demonstrating the potential of *in-situ* reflectance analysis, handheld instruments do not address the necessary automated data collection sought in an operational farm scenario. This challenge, however, can be surpassed through the use of UAVs and imaging systems (i.e., multispectral cameras) which are becoming ubiquitous in precision agriculture practices (Aasen et al., 2018).

Ideally, UAVs can fulfill an observational gap, providing the

* Corresponding author.

E-mail address: gustavo.togeirodealckmin@missouri.edu (G. Togeirode Alckmin).

<https://doi.org/10.1016/j.compag.2021.106574>

Received 19 January 2021; Received in revised form 13 November 2021; Accepted 15 November 2021

Available online 30 December 2021

0168-1699/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

necessary spatial, spectral and temporal resolution for farm operations, translating the accuracy achieved by handheld point-based methods to farm-scale aerial mapping. The difference in measured reflectance values due to data acquisition scale (i.e., from ground to low-level flight) has been shown negligible (Burkart et al., 2014), yet subject to sensor quality and design as well as environmental illumination conditions and appropriate radiometric calibration workflow (Olsson et al., 2021; Suomalainen et al., 2021). 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, the most popular UAV imaging systems (i.e., multi-camera arrays as per Aasen et al. (2018)) are limited to a small number of spectral bands and often rely on the use of *Vegetation Indices (VIs)* as enhanced predictors. Many of these indices, however, are prone to saturation, displaying an asymptotic response to high levels of biomass (Tucker, 1977), limiting model performance. Yet, saturation and poor model fit have been shown to be reduced when using a small subset of VIs and non-parametric regression models (Togeiro de Alckmin et al., 2021), further improving the potential and competitive advantage of UAVs imaging systems.

The technical widespread adoption of UAV imaging systems is limited due to data size and processing power requirements, leading to onerous scalability to large areas, as well as an absent automated pipeline from raw data to end-products (e.g., biomass maps). These constraints require predictive algorithms which can be quickly deployed for the generation of end-products from radiometrically corrected orthomosaics, either in a cloud-based system or on a personal computer. Both of these research-gaps (i.e., radiometrically accurate orthomosaics and scalable robust predictive models) are still challenging and are under active research (Pranga et al., 2021; Suomalainen et al., 2021).

From a sensing perspective, spectral data quality is subject to sensor design, data collection protocol and image processing from raw data to radiometrically corrected orthomosaics (Poncet et al., 2019). Recently, distinct studies have reported systematic measurement errors introduced by sensor design (Barker et al., 2020; Mamaghani and Salvaggio, 2019) and imagery processing workflow (Suomalainen et al., 2021). Consequently, despite a substantial demand, end-user products such as quantitative biophysical maps, which are dependent on radiometric accuracy, are not (yet) in place. Inevitably, to ensure a quantitative solution, imaging systems must provide accurate radiometric data for reliable deployment of predictive models (Fawcett and Anderson, 2019).

From a modeling perspective, the use of interpretable models, rather than black-boxes, should allow end-users to assess a key management parameter (e.g., biomass) while simultaneously indicating the drivers for that estimation (e.g., chlorophyll concentration or leaf-area index) through links between spectral predictors and biophysical inputs (Roscher et al., 2020). Such type of information may be employed towards model diagnosis, establishing under which conditions (e.g., LAI ≥ 5) models tend to under perform, and in which conditions specific spectral predictors are of greater importance. In summary, the deployment of prediction algorithms should ideally conjugate fast computation, strong generalization ability and interpretability.

However, there is usually a trade-off between (i) accuracy, (ii) model complexity and (iii) interpretability (Carvalho et al., 2019). For such reason, 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. Given the broadband spectral features measured in the context of field spectroscopy (Baret and Guyot, 1991), 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 models for optimal fit.

Furthermore, the prediction performances of machine learning models are typically communicated via *k*-fold cross-validation estimates, possibly providing an optimistic assessment. A more far-reaching evaluation of performance explores spatial and temporal assessments

(Meyer et al., 2018), providing insights to end-users about performance under different circumstances.

In summary, the triad of sensing, modelling and validation poses additive challenges in the understanding of achievable benefits of this technology in the context of pasture biomass estimation. For such reason, the objectives of this study are to (i) assess the correlation between reflectance measurements of a commercially available UAV multispectral camera at low-level flight and reference 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.

2. Methods

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 × 7.5 m, with 0.35 m border at each side of the plot's longitudinal axis, Fig. 1 - (IV)), arranged as two rows by 15 columns. Plots were grouped in three main blocks (10 plots per block, Fig. 1 - (IV)).

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. Within the blocks, plots were 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.

For the 2017 data collection campaign, each block was split in two 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 (Fig. 1 - (III)).

All plots in the 2018 campaign were initially mown and fertilized in October 21st, 2018. In the next two subsequent data collection campaigns, a single block was harvested and mown (at 21 and 27 days regrowth, 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) (Fig. 1 - (III)).

Data collection campaigns consisted of three subsequent stages: aerial multispectral imagery acquisition, (2) proximal hyperspectral measurements, and (3) biomass determination. 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 - Île-de-France, France, Fig. 1 - (II) Instruments). Four collection campaigns were performed under clear-sky conditions and around solar noon: November 28th, 2017 and the 11th, 17th, and 24th, November of 2018 (Table 1).

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. In total, 720 sample-sites were measured (i.e., 180 per data collection campaign - Table 1).

Within each plot, six sample-sites were randomly chosen (Fig. 1 - (I))

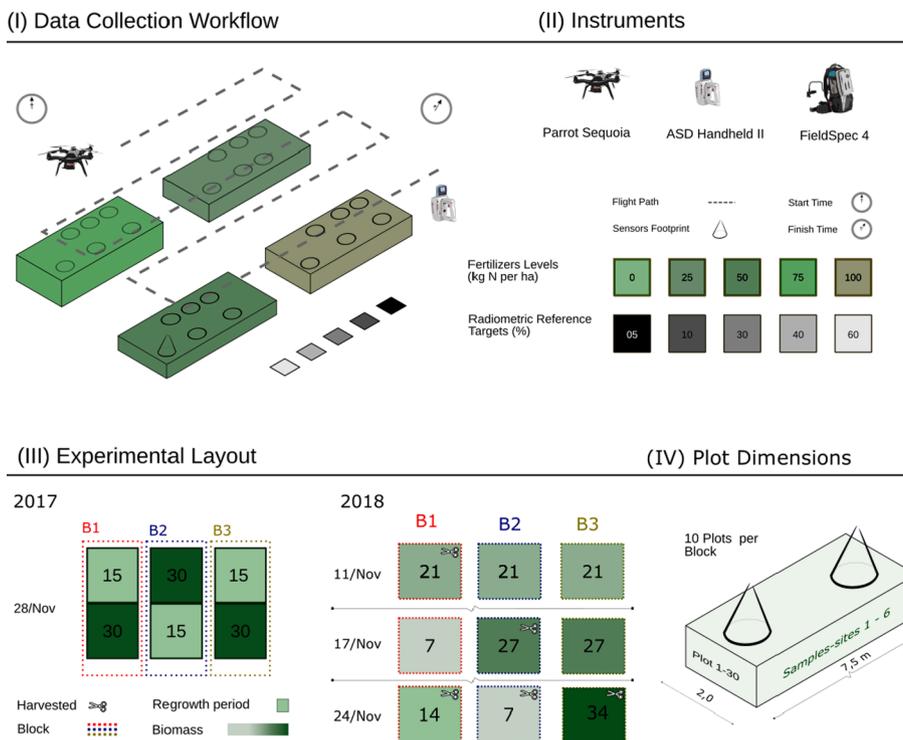


Fig. 1. General Overview of the Experimental Design. (I) Data Collection Workflow, (II) Instruments, (III) Experimental Layout for the 2017 and 2018 campaigns, and (IV) Plot Dimensions. In (III), the numbers within each block correspond to the number of growth days since last mowing event.

Table 1
Description of data collected in distinct campaigns.

Dates	Biomass Samples (n)	Spectral Samples (n)	Blocks Harvested	Sample Regrowth (days)
Nov-28-2017	180	180	B1, B2, B3	15 and 30
Nov-11-2018	60	180	B1	21
Nov-17-2018	60	180	B2	27
Nov-24-2018	180	180	B1, B2, B3	7, 14, 34

Data Collection Workflow). Before the start of the data collection campaign, markers were placed next to each sample-site to serve as visual references for the orthomosaics. Each sample-site was measured five times by the handheld spectrometers. The final sample-site 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. Proximal spectral measurements were taken from approximately one-meter height, thus, yielding a circular footprint equal to 0.15 m² (or 0.44 m diameter). Given the rigorous protocol, these spectral measurements are considered the benchmark to be compared against the multispectral camera.

Pasture biomass was mechanically defoliated above a residual height of 50 mm from the 0.15 m² spectrometer footprint. Harvested material was dried for a minimum of 48 h 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 (Table 1).

2.1.1. 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 h 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 × 40 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 × 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 ASD FieldSpec 4, and a contact probe. Measurements were taken 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).

2.1.2. 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 rather than the single Parrot panel). The orthomosaic was generated using the *MosaicMode* as *blending mode* and pixel size to 5 cm. This method is referred to as the *Pre-Calibration* method.

A second and subsequent correction was performed against the output of the *Pre-Calibration* method. This *Post-Calibration* method consisted in applying a linear function to the orthomosaic (i.e., empirical line method). This linear function was derived from the average reflectance of the known-reflectance targets as displayed in the orthomosaic and their (true) reflectance measured value in laboratory. Differently from the classic empirical line method, the linear function was applied after the proprietary Agisoft calibration was performed (i.e., modified Method “B” Poncet et al. (2019)) so to include the corrections derived from the irradiance sensor.

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).

2.3. Reference observations

To ensure comparability between modelling techniques, a *post hoc* analysis, the Dunn’s Multiple Comparison Test (`FSA::dunnTest`, Dinno (2015)) 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 (i.e., Parrot Sequoia). 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), as it is not made officially available yet disclosed by the manufacturer’s technical team.

For the multispectral imagery, polygon layers were generated to match the location of all handheld point measurements through a parametric script. 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 ($n = 720$, Table 1).

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, 2001). 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 (Fig. 2 - (III) Modelling).

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 (Fig. 2 - Modelling. Decision Trees and Cubist), what should considerably increase its computational speed.

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 (Fig. 2 - 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 significantly different from others, while grouping algorithms 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 l Index (LCI)* and *Modified Chlorophyll Absorption Ratio Index (MCARI)*, generated as per the equations and references available in `RStoolbox::spectralIndices`.

All the indices employed (Table 2) 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). Additionally, these indices were employed in Togeiro de Alckmin et al. (2020), allowing for model performance comparison between studies.

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 (Fig. 2 - (II) Validation Strategies).

3. Results

3.1. Reference observations

Biomass observations - For the complete DM data set, observations ranged from 132 to 4025 kg DM/ha, with an average value of 1356 and median 1125 kg DM/ha (Fig. 3 - 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 Fig. 3 - I.

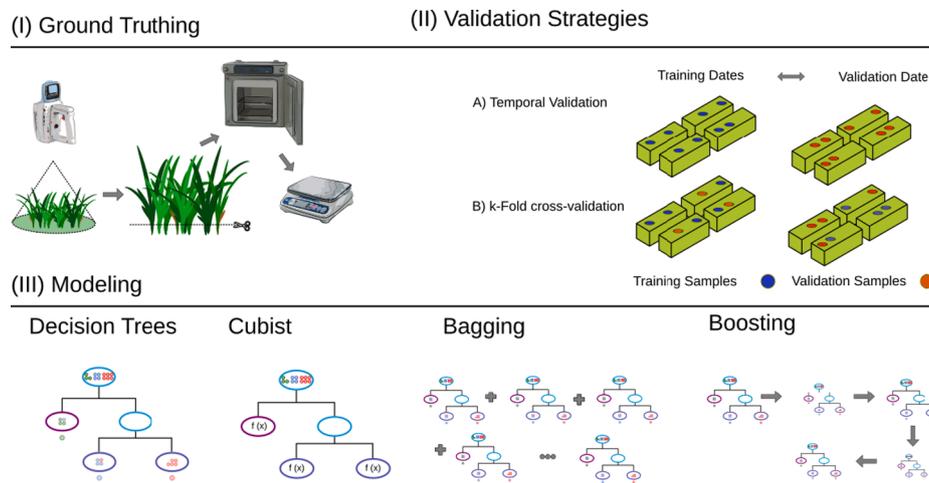


Fig. 2. General Overview of Sensing, Modelling and Validation. In (I) Ground Truthing, the sensor footprint is harvested, dried and weighed. In (II), both validation strategies are presented: (A) Temporal Validation, where a single date is set as validation set and (B) k-fold cross-validation, where no temporal constraint is imposed to the training and validation sets. In (III) Modeling, the structure of all modeling techniques is presented graphically.

Table 2
Vegetation indices employed for biomass modeling. Formula and Authorship are presented.

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)

observations of the first year within both Principal Components.

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 (Fig. 4 - 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 (Fig. 4 - Pre-Calibration). Such is made explicit 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 - 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 darker targets (i.g. Dark-Anthracite, Stone-Gray and Light-Gray) were employed as data points in the empirical line calibration (Fig. 4). For the remaining bands, all targets were used in the linear regression (Fig. 4).

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 (Fig. 4).

Radiometric Assessment - An assessment between the convolved handheld spectrometer data (i.e., the benchmark instrument) and the corrected orthomosaic ($n = 720$) is presented in the Table 3. 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 Fig. 5.

Particularly for the date of November 28th 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 17th 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 28th November, 2017.

Overall, all bands displayed a poor performance, presenting a

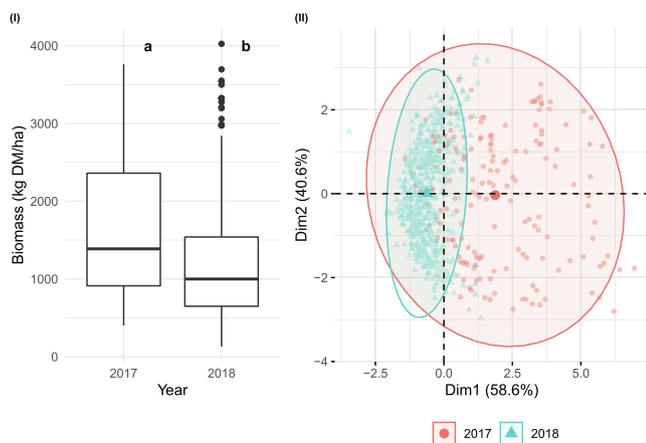


Fig. 3. Biomass and Spectral Observations - (I) Boxplot biomass values per year and (II) PCA biplot of ASD spectral measurements (benchmark for spectral measurements). In (I) a Dunn’s Test was performed and the observations were considered different (groups a and b) when aggregated by the factor exitYear. In (II), the ellipses represents 90% of the corresponding observations for the 2017 (red circles) and 2018 (blue triangles).

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. Fig. 3 - 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 (Fig. 3 - II). As a reduced feature space, the PCA biplot indicated the 2017 data displayed a wider spectral variability than 2018, encircling

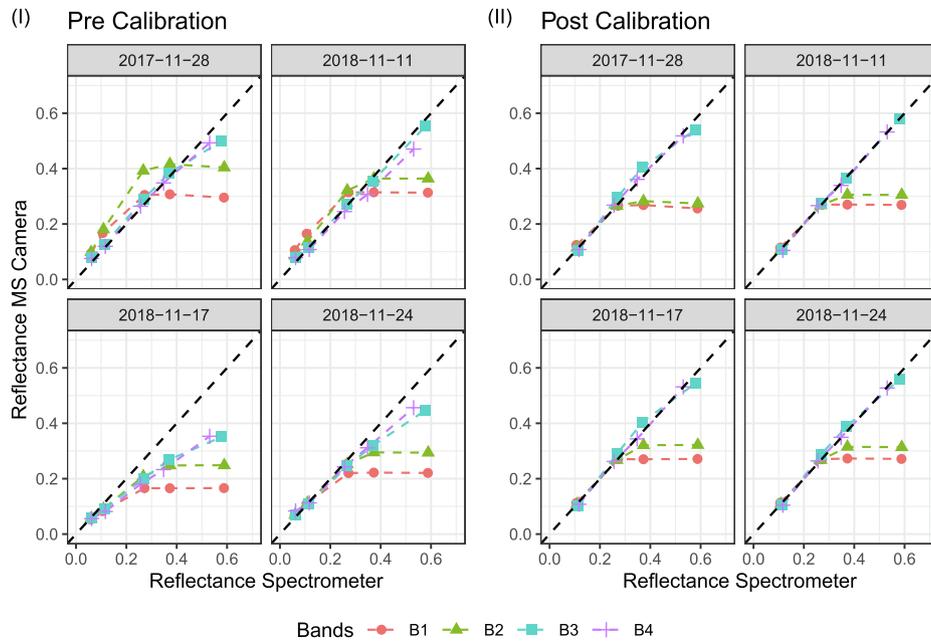


Fig. 4. Scatterplot of Multispectral Camera and Handheld Sensor Measurements of Reference Panels Spectral. Pre-Calibration refers to the radiometric calibration as per the manufacturer’s guidelines and extra reference panels (a variation of Method “B” described in Poncet et al. (2019)). Post Calibration refers to the sequential calibration employing a procedure analogous to the empirical line method.

Table 3
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

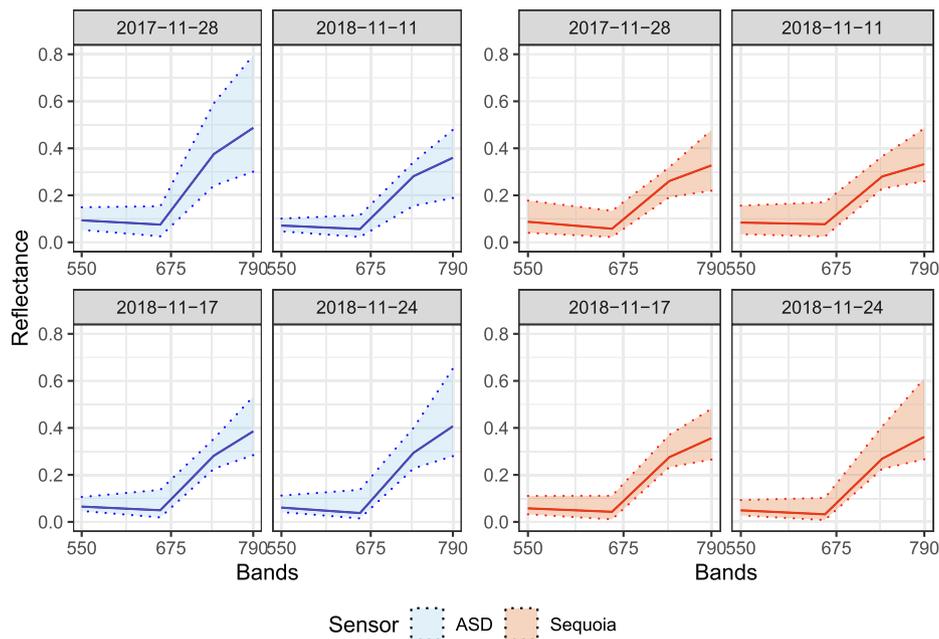


Fig. 5. Average, Maximum and Minimum Spectral Response per Date and per Instrument. ASD Fieldspec spectral response (left-side, blue). Parrot Sequoia spectral response (right-side, red). The coefficient of determination (R^2), RMSE and NRMSE, between both instrument is presented in Table 3.

median NRMSE equal to 19% Fig. 5.

3.3. Biomass modelling

The results of *k*-fold cross validation are presented in Fig. 6. Dunn’s Test were used to identify which specific means are significantly different from the others (i.e., different groups).

Multispectral camera - for the multispectral camera dataset (Fig. 6 - I), within the five different models, both Cubist and Random-Forest have the lowest and equivalent error distributions as per the results of Dunn’s test (Fig. 6 - 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 a 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 (Fig. 6 - 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 4), 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

Table 4

Model sizes and Prediction Time (based on the raster file from November 28th, 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

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 (Fig. 7 - 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

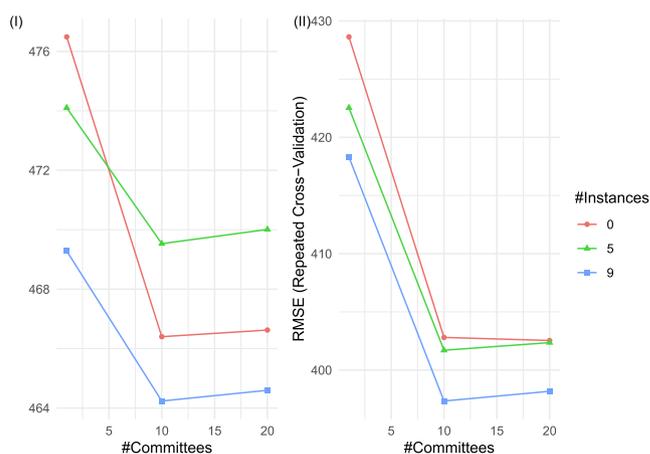


Fig. 7. Hyperparameter tuning for the Cubist regression algorithm. Sub-figures I and II corresponding to the model tuning using the handheld and the multispectral camera, respectively.

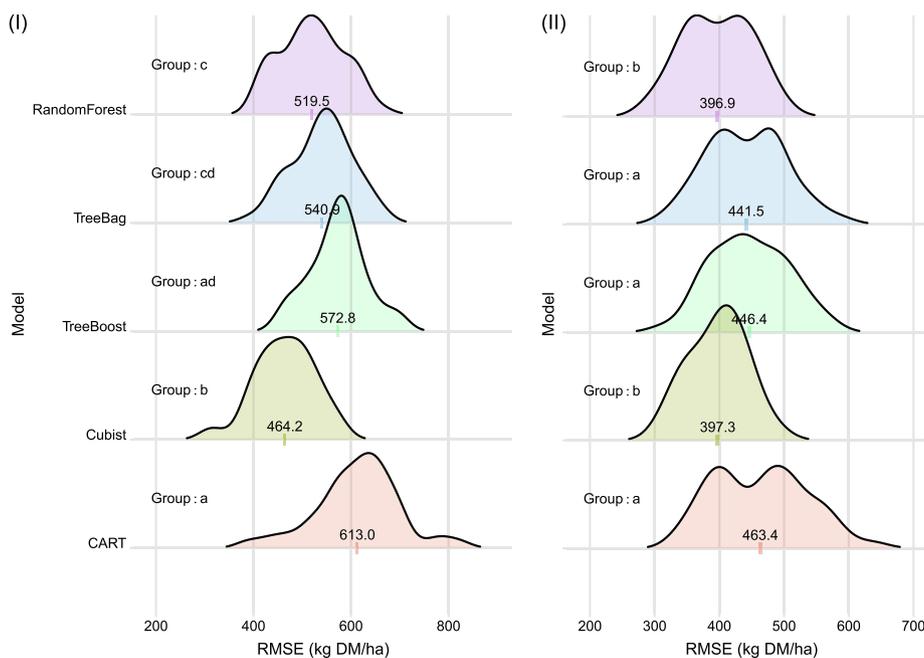


Fig. 6. Root Mean Square Error (RMSE) distribution for *k*-fold cross validation: (I) error-distribution for using a handheld spectrometer; (II) error-distribution using the multispectral camera data. A Dunn’s test (“Group”) was applied for each corresponding sensing instrument to determine whether error distributions were equivalent. Models size and processing times are presented in Table 4.

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, Fig. 7 - 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 (Fig. 8). For the purpose of brevity, only the Figure concerning the multispectral camera is presented, while the handheld is provided as an 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 (Fig. 9). 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).

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 - Fig. 9).

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 (Fig. 9). 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.

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 (Fig. 4). Consequently, the camera is limited to spectral observations of vegetation, for which most targets are within such range (i.e., reflectance below 20–30%). In summary, the reflectance measurements provided by the multispectral camera were not compatible to the reference handheld instrument.

Additionally, although our experimental design did not aim to measure it, the analysis (Fig. 4) indicates a lack of dynamic range of the multispectral camera in comparison to the handheld sensor (Fig. 5). 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 3). Finally, as the median NRMSE (19%), shows that the multispectral camera has a poor performance in its operational range when assessed against the handheld instruments.

Although the *Post-Calibration* method 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 an 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, as thoroughly discussed in Suomalainen et al. (2021) and Olsson et al. (2021).

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 its 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 Fig. 8, these models are far from complex and are a small number of linear models embedded in a

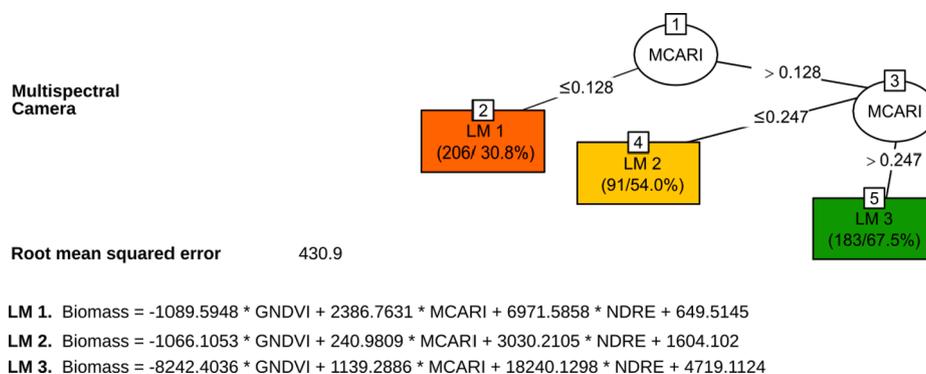


Fig. 8. Cubist Regression (multispectral camera and *k*-fold cross validation strategy). Terminal leaves present both the number of observation and coefficient of determination for its respective linear model (LM).

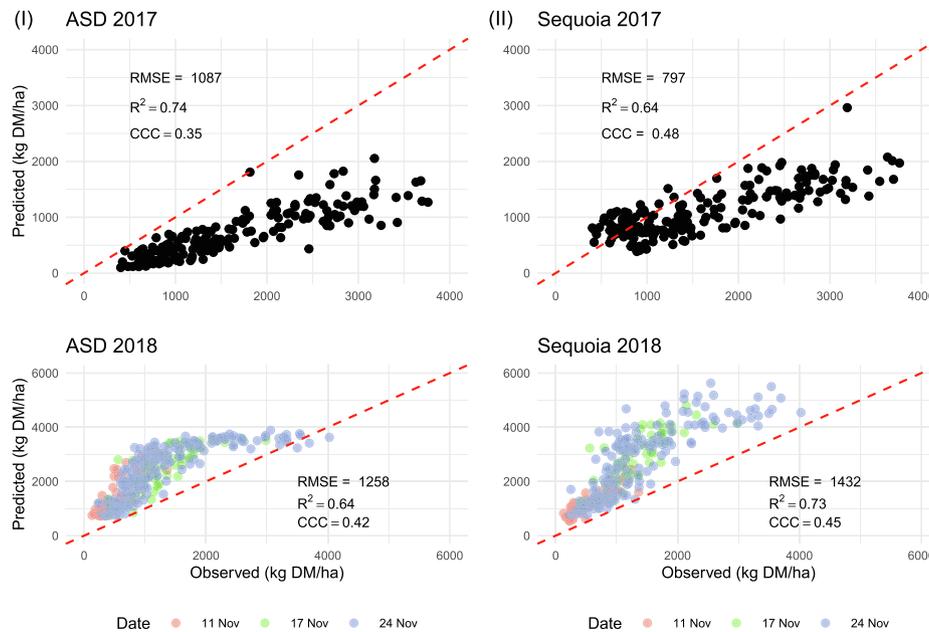


Fig. 9. Predictions for temporal validation strategy. Scatterplots on the left and right side, present the validation results for the handheld and camera data, respectively. Accordingly, the top and bottom row refer to the years used as validation sets, either 2017 or 2018, respectively.

CART algorithm, consequently, its deployment should be optimal in an operational scenario and comparable to standard linear models. 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 Fig. 8, 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 (such as Lasso, Ridge Regression and Elastic Nets), as high values of model coefficients indicated overfitting or multicollinearity, which would further improve the model's interpretability. These recommendations can be implemented in future studies and result in a more interpretable models, which are consequently more likely to be understood and adopted by the larger precision agriculture community.

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 Fig. 3. 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 an 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 manufacturers of multi-camera arrays, 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.

Yet, the saturation of the Green and Red Bands, may have been intentional, allowing a higher radiometric sensitivity (i.e., wider-dynamic range) for reflectance levels which are commonly found for vegetation (i.e., from 0% to 25% reflectance). 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) and more broadly in Suomalainen et al. (2021).

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 (Fig. 3 - (II)). Alternatively, the k-

fold cross-validation illustrates achievable accuracies and similar results as previous studies (Karunaratne et al., 2019; Pranga et al., 2021; Togeiro de Alckmin et al., 2020) as well as improvement from the results presented in Michez et al. (2019), where the authors employed the same multispectral camera for pasture biomass assessment. This RMSE range (300–400 kg DM per ha) is admissible for rotational grazing where pre-grazing biomass range from 1,500–2,500 kg DM per ha (Roca-Fernandez et al., 2011).

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. Conclusion

This study successfully compared matching spectral measurements of observations ($n = 720$), 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 ($n = 480$, biomass observations) were similar, with an average RMSE ranging from approximately 400 to 460 kg DM/ha for camera and handheld sensors in, 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. Although this particular camera has not shown a satisfactory process, superior methods have been proposed (Suomalainen et al., 2021), allowing for a direct radiometric calibration suitable for an operational farm scenario.

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. Further improvements were discussed to enhance its interpretability, although the algorithm output is significant improvement more interpretable than black-box models such as the popular Random-Forest algorithm.

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., 2021) and are surmountable. 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.

Funding

This research received no funding apart from those related to the Dairy on PAR action (Dairy Tasmania).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by Dairy Australia, through the Dairy on PAR action (Dairy Tasmania). In addition, the authors would like to express their gratitude to ASD Malvern PanAnalytical, which provided the FieldSpec 4 through the Goetz Instrument Program. In addition, the main author would like to express its gratitude for the instructive conversations with Dr. José Lucas Safannelli about R programming.

References

- Aasen, H., Honkavaara, E., Lucieer, A., Zarco-Tejada, P., 2018. Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures, and Data Correction Workflows. *Remote Sens.* 10, 1091. <https://doi.org/10.3390/rs10071091>.
- Agisoft, 2020. Agisoft Metashape User Manual.
- Baret, F., Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35, 161–173. [https://doi.org/10.1016/0034-4257\(91\)90009-U](https://doi.org/10.1016/0034-4257(91)90009-U).
- Barker, J.B., Woldt, W.E., Wardlow, B.D., Neale, C.M.U., Maguire, M.S., Leavitt, B.C., Heeren, D.M., 2020. Calibration of a common shortwave multispectral camera system for quantitative agricultural applications. *Precision Agric.* 21, 922–935. <https://doi.org/10.1007/s11119-019-09701-6>.
- Breiman, L., 1994. Bagging Predictors.
- Breiman, L., 2001. Random Forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 2017. Classification and Regression Trees, The wadsworth and brooks-cole statistics-probability series. Routledge, Boca Raton. <https://doi.org/10.1201/9781315139470>.
- Burkart, A., Cogliati, S., Schickling, A., Rascher, U., 2014. A Novel UAV-Based Ultra-Light Weight Spectrometer for Field Spectroscopy. *IEEE Sens. J.* 14, 62–67. <https://doi.org/10.1109/JSEN.2013.2279720>.
- Carvalho, D.V., Pereira, E.M., Cardoso, J.S., 2019. Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics* 8, 832. <https://doi.org/10.3390/electronics8080832>.
- Daughtry, C., 2000. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* 74, 229–239. [https://doi.org/10.1016/S0034-4257\(00\)00113-9](https://doi.org/10.1016/S0034-4257(00)00113-9).
- De Rosa, D., Basso, B., Fasiolo, M., Friedl, J., Fulkerson, B., Grace, P.R., Rowlings, D.W., 2021. Predicting pasture biomass using a statistical model and machine learning algorithm implemented with remotely sensed imagery. *Comput. Electron. Agric.* 180, 105880. <https://doi.org/10.1016/j.compag.2020.105880>.
- Dinno, A., 2015. Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. *Stata J.* 15, 292–300. <https://doi.org/10.1177/1536867x1501500117>.
- Fallet, C., Domenzain, L.M., 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: Messinger, D.W., Velez-Reyes, M. (Eds.), *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV*. International Society for Optics; Photonics; SPIE, p. 42. <https://doi.org/10.1117/12.2304334>.
- Fawcett, D., Anderson, K., 2019. Investigating impacts of calibration methodology and irradiance variations on lightweight drone-based sensor derived surface reflectance products 13. <https://doi.org/10.1117/12.2533106>.
- Franzini, M., Ronchetti, G., Sona, G., Casella, V., 2019. Geometric and Radiometric Consistency of Parrot Sequoia Multispectral Imagery for Precision Agriculture Applications. *Appl. Sci.* 9, 5314. <https://doi.org/10.3390/app9245314>.
- Fraser, B.T., Congalton, R.G., 2018. Issues in Unmanned Aerial Systems (UAS) data collection of complex forest environments. *Remote Sens.* 10 <https://doi.org/10.3390/rs10060908>.
- Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58, 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7).
- Gitelson, A., Merzlyak, M.N., 1994. Spectral Reflectance Changes Associated with Autumn Senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *J. Plant Physiol.* 143, 286–292. [https://doi.org/10.1016/S0176-1617\(11\)81633-0](https://doi.org/10.1016/S0176-1617(11)81633-0).
- Haas, R.H., Deering, D.W., Rouse, J.W., Schell, J.A., 1975. Monitoring vegetation conditions from Landsat for use in range management. In: *NASA Earth Resources Survey Symposium Proc.*, Houston, Texas, pp. 43–52.
- Karunaratne, S., Morse-mcnabb, E., Thomson, A., Stayches, D., Jacobs, J., 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, pp. 6–9.

- Lin, L.I.-K., 1989. A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics* 45, 255. <https://doi.org/10.2307/2532051>.
- Lu, H., Fan, T., Ghimire, P., Deng, L., 2020. Experimental evaluation and consistency comparison of UAV multispectral minisensors. *Remote Sens.* 12, 1–19. <https://doi.org/10.3390/RS12162542>.
- Mac Arthur, A.A., MacLellan, C., Malthus, T.J., 2007. The implications of non-uniformity in fields-of-view of commonly used field spectroradiometers. In: 2007 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Barcelona, pp. 2890–2893. <https://doi.org/10.1109/IGARSS.2007.4423447>.
- Mamaghani, Salvaggio, 2019. Multispectral Sensor Calibration and Characterization for sUAS Remote Sensing. *Sensors* 19, 4453. <https://doi.org/10.3390/s19204453>.
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., Naus, T., 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environ. Model. Softw.* 101, 1–9. <https://doi.org/10.1016/j.envsoft.2017.12.001>.
- Michez, A., Lejeune, P., Bauwens, S., Herinaina, A., Blaise, Y., Castro Muñoz, E., Lebeau, F., Bindelle, J., Lalaina Herinaina, A.A., Blaise, Y., Muñoz, E.C., Lebeau, F., Bindelle, J., 2019. Mapping and monitoring of biomass and grazing in pasture with an unmanned aerial system. *Remote Sens.* 11, 473. <https://doi.org/10.3390/rs11050473>.
- Mutanga, O., Skidmore, A.K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* 25, 3999–4014. <https://doi.org/10.1080/01431160310001654923>.
- Olsson, P.O., Vivekar, A., Adler, K., Garcia Millan, V.E., Koc, A., Alamrani, M., Eklundh, L., 2021. Radiometric correction of multispectral uas images: Evaluating the accuracy of the parrot sequoia camera and sunshine sensor. *Remote Sens.* 13, 1–26. <https://doi.org/10.3390/rs13040577>.
- Opitz, D., Maclin, R., 1999. Popular Ensemble Methods: An Empirical Study. *J. Artif. Intell. Res.* 11, 169–198. <https://doi.org/10.1613/jair.614>.
- Poncet, A.M., Knappenberger, T., Brodbeck, C., Fogle, M., Shaw, J.N., Ortiz, B.V., 2019. Multispectral UAS Data Accuracy for Different Radiometric Calibration Methods. *Remote Sens.* 11, 1917. <https://doi.org/10.3390/rs11161917>.
- Pranga, J., Borra-Serrano, I., Aper, J., De Swaef, T., Ghesquiere, A., Quataert, P., Roldán-Ruiz, I., Janssens, I.A., Ruysschaert, G., Lootens, P., 2021. Improving Accuracy of Herbage Yield Predictions in Perennial Ryegrass with UAV-Based Structural and Spectral Data Fusion and Machine Learning. *Remote Sens.* 13, 3459. <https://doi.org/10.3390/rs13173459>.
- Pricope, N.G., Mapes, K.L., Woodward, K.D., Olsen, S.F., Baxley, J.B., 2019. Multi-Sensor Assessment of the Effects of Varying Processing Parameters on UAS Product Accuracy and Quality. *Drones* 3, 63. <https://doi.org/10.3390/drones3030063>.
- Quinlan, J.R., 1992. Learning with continuous classes, *Australian Joint Conference on Artif. Intell.* 343–348.
- Quinlan, J.R., 1993. Combining Instance-Based and Model-Based Learning. In: *Machine Learning Proceedings 1993*. Elsevier, pp. 236–243. <https://doi.org/10.1016/B978-1-55860-307-3.50037-X>.
- Roca-Fernandez, A.I., O'Donovan, M.A., Curran, J., Gonzalez-Rodriguez, A., 2011. Effect of pre-grazing herbage mass and daily herbage allowance on perennial ryegrass swards structure, pasture dry matter intake and milk performance of Holstein-Friesian dairy cows. *Spanish J. Agric. Res.* 9, 86. <https://doi.org/10.5424/sjar/20110901-126-10>.
- Rohantgi, A., 2020. WebPlotDigitizer.
- Roscher, R., Bohn, B., Duarte, M.F., Garcke, J., 2020. Explain it to me-facing remote sensing challenges in the bio-and geosciences with explainable machine learning. *ISPRS Ann. Photogram. Remote Sens. Spatial Inform. Sci.* 5, 817–824. <https://doi.org/10.5194/isprs-Annals-V-3-2020-817-2020>.
- Rouse, J.W., Hass, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the great plains with ERTS. In: *Third Earth Resources Technology Satellite (ERTS) symposium*, vol. 1, pp. 309–317.
- Schapiro, R.E., 1990. The Strength of Weak Learnability 227, 197–227.
- Stone, M., 1974. Cross-Validatory Choice and Assessment of Statistical Predictions. *J. Roy. Stat. Soc. Ser. B (Methodol.)* 36, 111–147.
- Suomalainen, J., Hakala, T., de Oliveira, R.A., Markelin, L., Viljanen, N., Näsi, R., Honkavaara, E., 2018. A novel tilt correction technique for irradiance sensors and spectrometers on-board unmanned aerial vehicles. *Remote Sens.* 10, 1–18. <https://doi.org/10.3390/rs10122068>.
- Suomalainen, J., Oliveira, R.A., Hakala, T., Koivumäki, N., Markelin, L., Näsi, R., Honkavaara, E., 2021. Direct reflectance transformation methodology for drone-based hyperspectral imaging. *Remote Sens. Environ.* 266, 112691. <https://doi.org/10.1016/j.rse.2021.112691>.
- Togeiro de Alckmin, G., Kooistra, L., Rawnley, R., de Bruin, S., Lucieer, A., 2020. Retrieval of hyperspectral information from multispectral data for perennial ryegrass biomass estimation. *Sensors (Switzerland)* 20, 1–20. <https://doi.org/10.3390/s20247192>.
- Togeiro de Alckmin, G., Kooistra, L., Rawnley, R., Lucieer, A., 2021. Comparing methods to estimate perennial ryegrass biomass: canopy height and spectral vegetation indices. *Precision Agric.* 22, 205–225. <https://doi.org/10.1007/s11119-020-09737-z>.
- Tucker, C.J., 1977. Asymptotic nature of grass canopy spectral reflectance. *Appl. Opt.* 16, 1151. <https://doi.org/10.1364/AO.16.001151>.
- Wilkinson, J.M., Lee, M.R.F., Rivero, M.J., Chamberlain, A.T., 2020. Some challenges and opportunities for grazing dairy cows on temperate pastures. *Grass Forage Sci.* 75, 1–17. <https://doi.org/10.1111/gfs.12458>.
- Zebarth, B.J., Younie, M., Paul, J.W., Bittman, S., 2002. Evaluation of leaf chlorophyll index for making fertilizer nitrogen recommendations for silage corn in a high fertility environment. *Commun. Soil Sci. Plant Anal.* 33, 665–684. <https://doi.org/10.1081/CSS-120003058>.