



Article

Broadacre Crop Yield Estimation Using Imaging Spectroscopy from Unmanned Aerial Systems (UAS): A Field-Based Case Study with Snap Bean

Amirhossein Hassanzadeh ^{1,*} , Fei Zhang ¹ , Jan van Aardt ¹ , Sean P. Murphy ² and Sarah J. Pethybridge ²

¹ Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, NY 14623, USA; fz4070@rit.edu (F.Z.); jvacis@rit.edu (J.v.A.)

² Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell AgriTech at The New York State Agricultural Experiment Station, Cornell University, Geneva, NY 14456, USA; sm2244@cornell.edu (S.P.M.); sjp277@cornell.edu (S.J.P.)

* Correspondence: ah7557@rit.edu

Abstract: Accurate, precise, and timely estimation of crop yield is key to a grower's ability to proactively manage crop growth and predict harvest logistics. Such yield predictions typically are based on multi-parametric models and in-situ sampling. Here we investigate the extension of a greenhouse study, to low-altitude unmanned aerial systems (UAS). Our principal objective was to investigate snap bean crop (*Phaseolus vulgaris*) yield using imaging spectroscopy (hyperspectral imaging) in the visible to near-infrared (VNIR; 400–1000 nm) region via UAS. We aimed to solve the problem of crop yield modelling by identifying spectral features explaining yield and evaluating the best time period for accurate yield prediction, early in time. We introduced a Python library, named Jostar, for spectral feature selection. Embedded in Jostar, we proposed a new ranking method for selected features that reaches an agreement between multiple optimization models. Moreover, we implemented a well-known denoising algorithm for the spectral data used in this study. This study benefited from two years of remotely sensed data, captured at multiple instances over the summers of 2019 and 2020, with 24 plots and 18 plots, respectively. Two harvest stage models, early and late harvest, were assessed at two different locations in upstate New York, USA. Six varieties of snap bean were quantified using two components of yield, pod weight and seed length. We used two different vegetation detection algorithms. the Red-Edge Normalized Difference Vegetation Index (RENDVI) and Spectral Angle Mapper (SAM), to subset the fields into vegetation vs. non-vegetation pixels. Partial least squares regression (PLSR) was used as the regression model. Among nine different optimization models embedded in Jostar, we selected the Genetic Algorithm (GA), Ant Colony Optimization (ACO), Simulated Annealing (SA), and Particle Swarm Optimization (PSO) and their resulting joint ranking. The findings show that pod weight can be explained with a high coefficient of determination ($R^2 = 0.78\text{--}0.93$) and low root-mean-square error ($RMSE = 940\text{--}1369$ kg/ha) for two years of data. Seed length yield assessment resulted in higher accuracies ($R^2 = 0.83\text{--}0.98$) and lower errors ($RMSE = 4.245\text{--}6.018$ mm). Among optimization models used, ACO and SA outperformed others and the SAM vegetation detection approach showed improved results when compared to the RENDVI approach when dense canopies were being examined. Wavelengths at 450, 500, 520, 650, 700, and 760 nm, were identified in almost all data sets and harvest stage models used. The period between 44–55 days after planting (DAP) the optimal time period for yield assessment. Future work should involve transferring the learned concepts to a multispectral system, for eventual operational use; further attention should also be paid to seed length as a ground truth data collection technique, since this yield indicator is far more rapid and straightforward.

Keywords: feature selection; hyperspectral imaging; machine learning; snap bean; unmanned aerial vehicle; yield modelling



Citation: Hassanzadeh, A.; Zhang, F.; van Aardt, J.; Murphy, S.P.; Pethybridge, S.J. Broadacre Crop Yield Estimation Using Imaging Spectroscopy from Unmanned Aerial Systems (UAS): A Field-Based Case Study with Snap Bean. *Remote Sens.* **2021**, *13*, 3241. <https://doi.org/10.3390/rs13163241>

Academic Editor: Brigitte Leblon

Received: 7 July 2021

Accepted: 11 August 2021

Published: 15 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Agriculture and related industries contributed more than \$1 trillion to the United States' gross domestic product (GDP) in 2019 [1]. However, the agricultural industry is a sector vulnerable to climate change, market factors, and within-season weather variability. Inventory management in this sector, generally referred to as yield evaluation, is different from other sectors, since crops require time for growth, and estimating "how much" and "how many" are dependent on the physiology of living organisms. Nowadays, growers rely on "expert guesses", via multi-parameter models or in-situ assessments, which often provide unreliable estimates. These inaccuracies could heavily impact not only the well-being of a business, but on a larger scale, the economy of a region or a country. Growers also need to accurately set the price of the crop with buyers prior to harvest, gauged based on the predicted yield estimate. If the actual final yield is more than the estimated yield, there is a chance the farmer would have to sell the remainder for a discounted price to compensate for the unplanned surplus. Conversely, if the actual final yield is less than the estimated yield, the business's net profit is damaged because the unexpected shortage of yield means they cannot sell as much as they projected. In both scenarios, buyers and customers could become reluctant to continue doing business with the growers because they are not receiving what they were promised. Accurate knowledge of the final yield thus is important to both growers and the market to manage price.

Precision agriculture (PA) could provide an answer to this need by maximizing profit and minimizing expenses and resources use. PA presents a systematic optimization technique and, when tied to remote sensing (RS), can provide non-destructive, rapid, and affordable information about the target being studied [2,3]. RS modalities include in-situ, airborne, or spaceborne systems, accommodating various sensors from RGB (color), multi-spectral, hyperspectral, and light detection and ranging (LiDAR), each suited to a specific application and objective [4]. These systems can be used for various crop assessments, such as nutrient levels [5,6], water stress [7], weed detection [8], soil properties [9], crop growth [10], and crop biomass or yield [9,11–15]. However, to assess such crop characteristics from a remote platform, researchers often rely on hyperspectral sensing systems.

Hyperspectral systems are at the forefront of PA, given their ability to collect hundreds of contiguous narrow spectral bands over a specified range in the electromagnetic spectrum [16,17]. This enables a targeted approach to develop more operational, cost-effective multispectral solutions, distilled from the hyperspectral research. One feature of crop production is that spectral signatures of the canopy change as crops mature [4]. This variation could be valuable information to monitor as it enables an assessment of molecular absorption features, foliar chemical characteristics, and particle scattering of the crop [18–20]. This change is noticeable in terms of amplitude, absorption features, and change in the *slope*, of the spectrum [4]. One caveat to the use of such systems, however, involves the need to address potential noise contained in narrow bandwidth, large spectral range systems [21].

The quality of the data collected by hyperspectral sensor is degraded by atmospheric effects, such as solar illumination and aerosol scattering, as well as instrument-based noise, such as photon noise, quantization noise, etc. [20,22,23]. Moreover, when a noisy signal is fed to a machine learning or a statistical model, it could be interpreted as valuable information, due to the inherent high variance. Hyperspectral noise reduction has been studied extensively in the literature and different methods have been proposed for signal noise reduction, such as state-of-the-art deep learning methods, e.g., auto-encoders [24–26], as well as Fourier transforms, derivatives, polynomial fitting methods [27–29], and wavelet transforms [30,31]. Among the mentioned approaches, the study of Chen et al. (2010) has attracted much attention in the past decade, where the variability in data is taken into account using principal component analysis (PCA), followed by noise reduction via wavelet shrinkage [32]. To add more detail, the original spectral domain is converted to principal component (PC) space and a specific number of PC's are selected as being important to retain, while the remaining components are denoised using discrete wavelet shrinkage. The denoised PC's are fused back with the original retained PC's, and finally converted

back to the original spectral domain. Compared to other methods, this approach is far more intuitive since the valuable information is detected based on the explained variability (i.e., eigenvalues) and then the noise is mitigated for PC's with smaller explained variability. However, the introduced approach performs well under the assumption that the proper number of PC's are retained, since important information could become lost if incorrect PC's are selected. Zwick et al. (1986) compared five different methods for determining the number of components to retain for PCA [33]. These rules were Bartlett's test (BART; [34]), Eigenvalues greater than 1 (K1; [35]), minimum average partial (MAP; [36]), the SCREE test [37], and parallel analysis [38]. The authors reported consistent high performance for all data sets employing the MAP rule, which is based on a matrix of partial correlations and the calculated squared partial correlation. Generating a spectral signal with reduced noise is a first crucial step in hyperspectral analysis; however, hyperspectral data comes with redundant amount of information that also needs to be addressed [39,40].

Hyperspectral data are known as rich, high-dimensional data that require specific approaches to analysis. There are two main aspects that one needs to consider when using hyperspectral data. First, a large number of bands typically are highly correlated, with the extent to which one variable is correlated with another being quantifiable via Pearson's correlation coefficient [41]. Feeding highly correlated variables to machine learning algorithms can cause overfitting [42]. Second, even after selecting uncorrelated bands, there remains a small number of bands that actually contribute to the objective function being studied. The process of selecting discriminating features is called feature selection, which is a pivotal step in multivariate analyses. Feature selection approaches fall under the three main categories of filter, wrapper, and embedded [43]. Filter methods determine feature importance (i.e., scoring) based on the independent variables and response variables in the data set (independent of the objective function), with examples such as Pearson's correlation coefficient, mutual information, etc. [44–46]. Filter methods have the drawback of not taking into account the influence of the combination of features (as a subset) on model performance. In contrast, wrapper algorithms are of an inductive nature and wrap around the objective function (dependent on the objective function), while taking into account the effect of features on each other as a subset. However, these methods are computationally more expensive, due to their iterative nature. Examples of this category include genetic algorithm (GA; [47]), ant colony optimization (ACO; [48]), particle swarm optimization (PSO; [49]), and sequential search [50]. Embedded methods include algorithms that identify a subset of features in the process of training, but with a lower computational cost than wrapper methods [46]. Examples in this category are regularization techniques [51]. Overall, hyperspectral imagery, from data acquisition and preprocessing, to band selection, requires attention in terms of processing and modelling steps, but past prediction successes have highlighted this modality's usefulness.

There have been numerous recent published studies focusing on yield estimation using UAS-based hyperspectral sensing, here we mainly focus on machine learning approaches. Li et al. (2020) used RGB and hyperspectral sensing (400–1000 nm) as a basis for yield modelling of potatoes. Their results show that fresh and dry biomass could be estimated with a high coefficient of determination ($R^2 = 0.9$) using random forests (RF) regression and RReliefF feature selection [14]. Tao et al. (2020) studied wheat yield estimation in China using hyperspectral imagery at three different stages of growth. Their approach included extracting spectral indices from the hyperspectral data, and they reported acceptable R^2 values of 0.72 and a root-mean-square error (RMSE) of 648 kg/ha using partial least squares regression (PLSR) at the flowering stage [52]. Feng et al. (2020) adopted an ensemble approach to model yield for alfalfa in Wisconsin, USA. Their approach included generating vegetation indices (VIs) and feeding these into RF, support vector regression (SVR), and K-nearest neighbors (KNN) models, while using the recursive feature elimination (RFE) feature selection method [15]. Their final models show an R^2 of 0.87 and an RMSE of 220 kg/ha using the ensemble approach. However, among various crops studied

in the literature for yield assessment, one crop lacks fundamental research regarding yield assessment, namely beans and specifically, snap bean.

Snap bean (*Phaseolus vulgaris*) is one of the largest sources of agricultural income, with a combined value of over \$400 million in 2016 for the USA [53]. Yuan et al. (2017) studied snap bean yield and growth as a function of nitrogen content. Their findings showed a high coefficient of determination ($R^2 = 0.90$) and a low error ($RMSE = 7.6\%$) [54]. Saleh et al. (2018) studied the effect of irrigation on yield and growth of snap bean crop for two different cultivars and identified 80% ET as optimal irrigation application [55]. Our previously published greenhouse study, using a handheld spectroradiometer in the VIS-SWIR (400–2500 nm), intended to model yield of snap bean crop (cv. Huntington), in a best-case scenario environment, at two harvest times (early and late). We showed that snap bean yield could be predicted, early in the growing season, most accurately 25–32 days prior to harvest with $RMSE$ values as low as 3 g/plant and R^2 values as high as 0.72. To the best of our knowledge, no studies have been conducted to evaluate the relationships between hyperspectral reflectance and yield components in snap bean in field conditions.

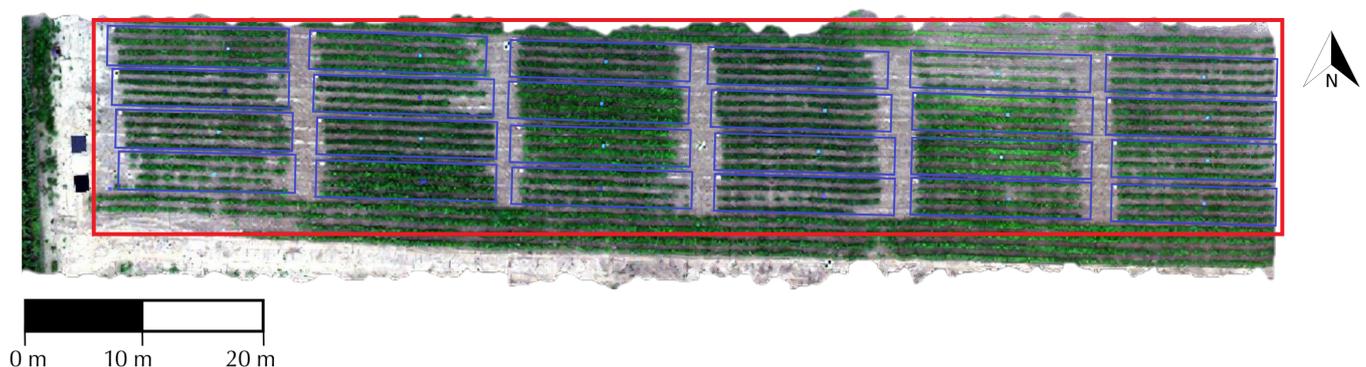
We hypothesize that yield modelling assessment of snap bean can be conducted using airborne hyperspectral systems, since plant yield has been shown to be a function of spectral responses, where spectral response reflects biophysical and physiological characteristics. Thus, the objectives of this study were to:

1. Assess yield prediction of snap bean, at various points during the growing season, using hyperspectral imagery and descriptive models;
2. Identify discriminating spectral features explaining yield; and
3. Evaluate the most accurate time (growth period) for yield prediction, prior to harvest.

2. Methods

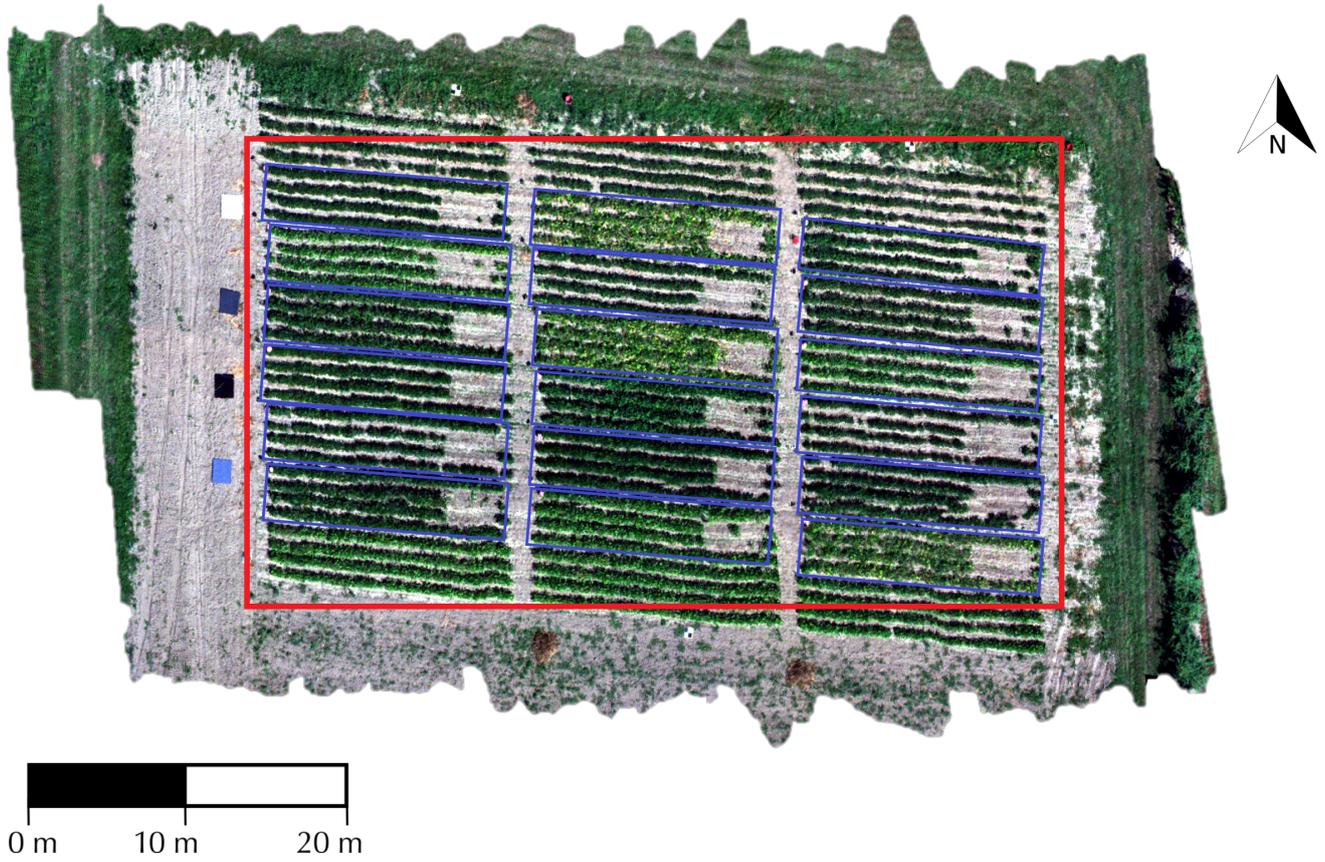
2.1. Study Area

The study took place at two different geographical locations during summer 2019 at Geneva, NY (42°49′53.0″N, 77°00′48.2″W) and summer 2020 at Seneca, NY (42°51′59.2″N, 77°01′45.8″W). Individual plots were identified (24 and 18 in 2019 and 2020, respectively; Figure 1). Six different cultivars of snap bean were chosen for a holistic evaluation across different snap bean varieties: Venture (large bean cultivar; L1), Huntington (large bean cultivar 2; L2), Colter (four sieve cultivar 1; F1), Cabot (four sieve cultivar 2; F2), Flavor-sweet (whole bean cultivar 1; W1), and Denver (whole bean cultivar 2; W2). These six cultivars were replicated four times and three times, for 2019 and 2020 data, respectively (Figure 2).



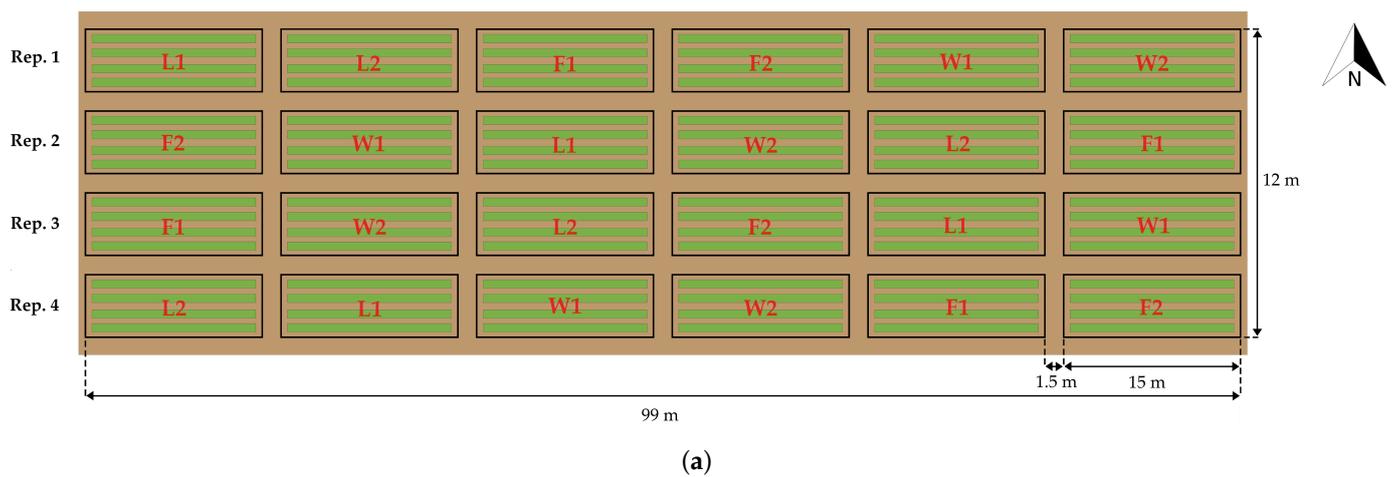
(a)

Figure 1. Cont.



(b)

Figure 1. RGB representation of the data and the designated plots for (a) 2019 and (b) 2020 experimental designs. Note the plots' bounds in blue.



(a)

Figure 2. Cont.

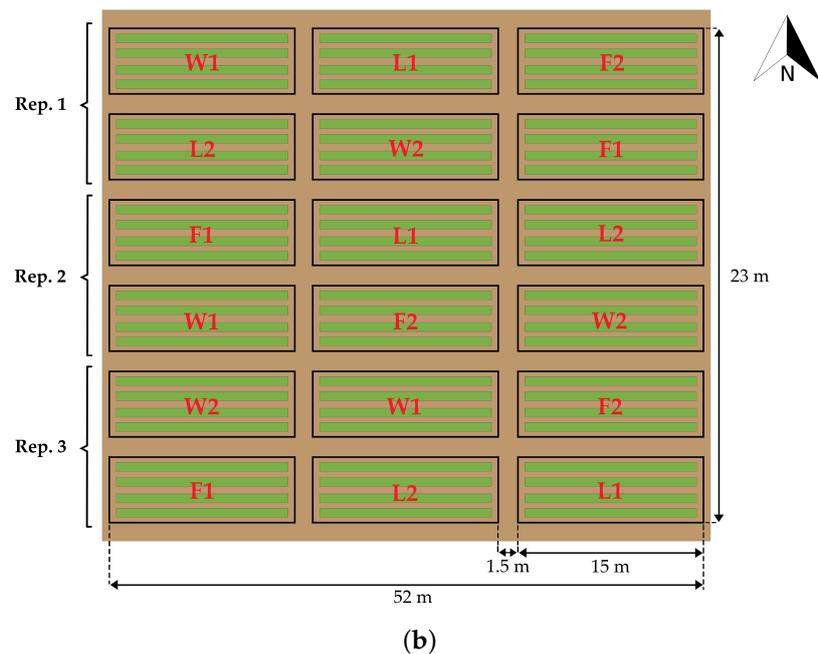


Figure 2. Outlined plots and matching cultivars along with field and plot design dimensions for (a) 2019 and (b) 2020 experimental designs. Listed here are the Venture (large cultivar variety 1; L1), Huntington (large cultivar variety 2; L2), Colter (four sieve cultivar 1; F1), Cabot (four sieve cultivar 2; F2), Flavor-sweet (whole bean cultivar 1; W1), and Denver (whole bean cultivar 2; W2). Schematics are not to scale.

2.2. Assessment of Plant Growth Characteristics

Each data set was considered an individual model and each plot was considered one experimental unit. Ground truth yield components (pod weight and seed length) were measured in 3-m sub-sampled areas within plots. Pod weight was weighed with the use of a balance (Model “ES6R”, OHAUS Corp., Parsippany, NJ, USA) with 0.002 kg precision. Seed length was measured using a maturity gauge (Seminis, Missouri, USA; Figure 3). The same gauge was used to classify pods into different sieve sizes for determining harvest timing (scheduling). Sieve size one through six were designated S1–S6, with sizes as follows: S1 < 5.8 mm; 5.8 mm < S2 < 7.5 mm; 7.5 mm < S3 < 8.5 mm; 8.5 mm < S4 < 9.7 mm; 9.7 mm < S5 < 10.9 mm; S6 > 10.9 mm. After pods were classified into sieve sizes, 10 seeds from 10 different pods from the largest sieve size were extracted and placed alongside each other on the indented part of the seed length gauge, where the measured length was recorded and reported as seed length. To assess the influence of harvest timing, two different harvest stages (early vs. late) were assessed. The early harvest model was determined when the majority of the Venture cultivar (L1) had reached S6, Huntington (L2) had reached S5, Colter (F1) had reached S4, Cabot (F2) had reached S4, Flavor-sweet (W1) had reached S2, and Denver (W2) had reached its maximum length, which is S1. Late harvest occurred two days after the early harvest for both 2019 and 2020 data sets.

Each plot contained four rows for both the 2019 and 2020 data. Seeds were planted using a Monosem precision planter (Monosem Inc., Edwardsville, KS, USA) at a density of 20–26 plants/m, depending on the cultivar, at a depth of approximately 3.8 cm. Fertilizer with an N:P:K ratio of 15:5:5 was only applied at planting at a rate of 335 kg/ha. No irrigation was applied to the trial.

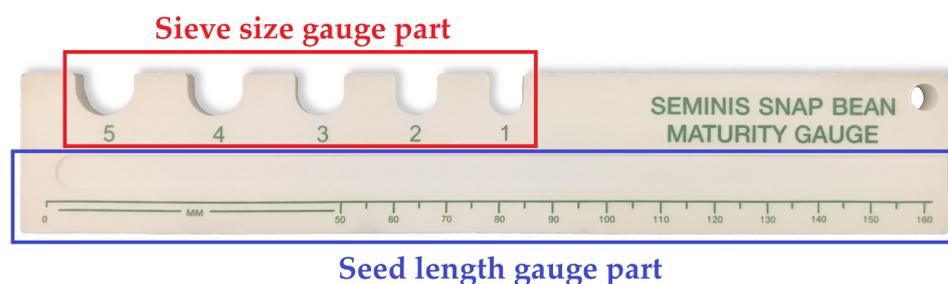


Figure 3. Maturity gauge used for determining seed length and pods' sieve sizes. On the upper part, pod sieve size is determined by sliding pods across the cavities. On the bottom part, ten seeds are put alongside each other and the total length is recorded.

2.3. Data Collection

Hyperspectral data were collected using an imaging spectrometer (“Nano Hyper-spec” model, Headwall Photonics, Fitchburg, Massachusetts, USA) in the VNIR domain (400–1000 nm), with 272 contiguous spectral bands. The hyperspectral sensor incorporates a sensor array of 640 pixels, and a full-width-half-maximum (FWHM) of 6 nm. The Nano sensor was mounted on a customized DJI Matrice-600 quadcopter (MX-1; [56]). Data were collected seven and eight times during the growth season, for 2019 and 2020 data, respectively (Table 1).

Table 1. Growth calendar of 2019 and 2020 experimental designs along with the captured ground sampling distance (GSD) of a snap bean field in Geneva, New York, USA. Please note that the first date is the sowing date and no data were collected.

Model	Date	Stage	Days after Planting (DAP)	Heat Unit (°) *	GSD (cm)
2019	06/27/2019	Sowing	0	0	N/A
	08/01/2019	Flowering	35	794	3
	08/05/2019	Flowering	39	865	3
	08/12/2019	Pod formation	46	997	1.5
	08/14/2019	Pod formation	48	1040	3
	08/16/2019	Pod formation	50	1074	3
	08/20/2019	Pod formation (early harvest)	54	1158	3
	08/22/2019	Pod formation (late harvest)	56	1205	3
2020	06/27/2020	Sowing	0	0	N/A
	07/28/2020	Budding	31	760	3
	07/31/2020	Flowering	34	822	3
	08/06/2020	Flowering	40	945	3
	08/10/2020	Pod formation	44	1022	3
	08/14/2020	Pod formation	48	1116	3
	08/21/2020	Pod formation	55	1239	3
	08/24/2020	Pod formation (early harvest)	58	1306	3
08/26/2020	Pod formation (late harvest)	60	1346	3	

* heat unit calculated with $T_{base} = 10$ °C.

2.4. Data Preprocessing

The data preprocessing section was comprised of three primary steps, which are calibration to reflectance, vegetation detection, and spectral denoising, as shown in the flowchart in Figure 4.

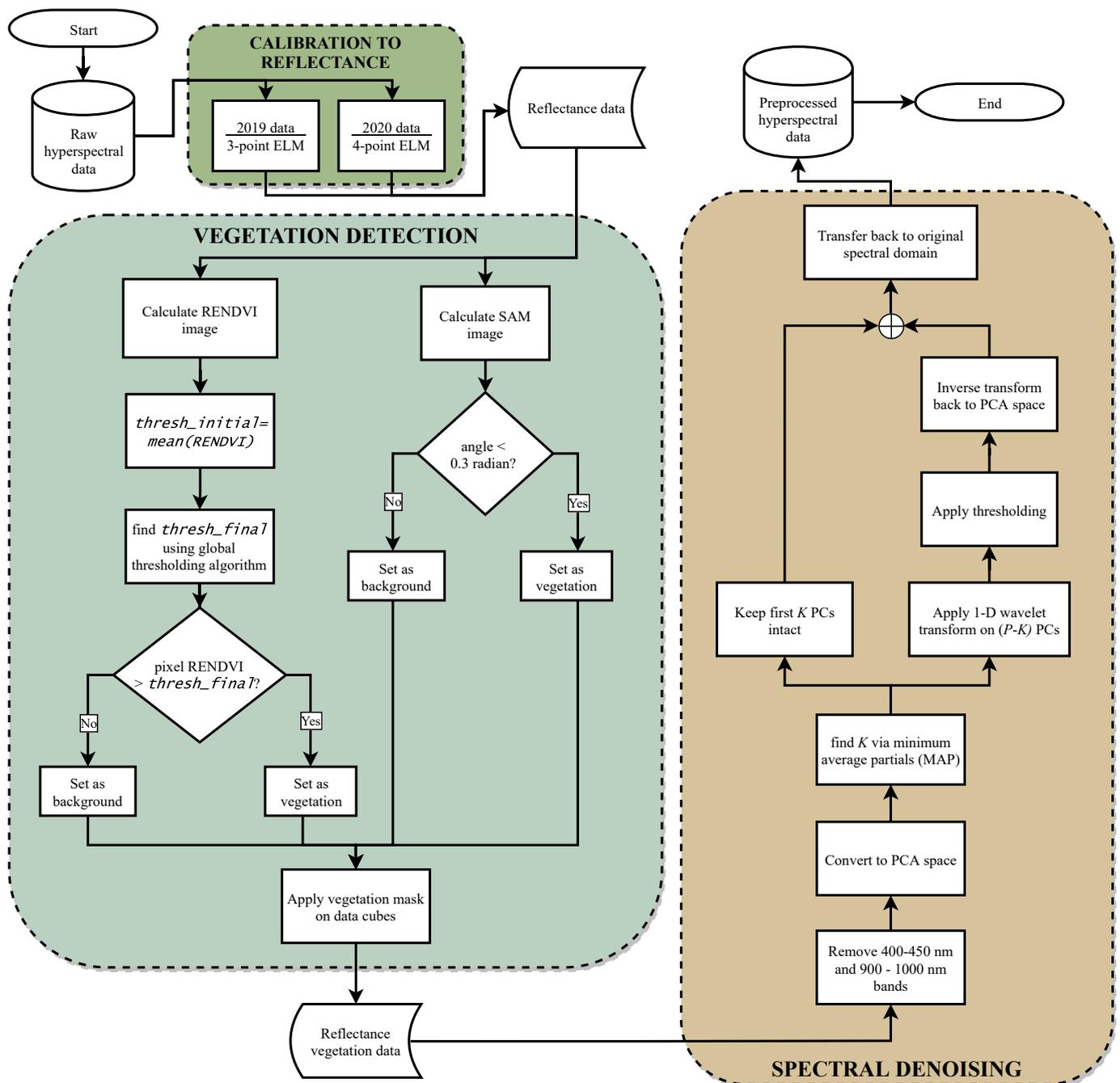


Figure 4. Flowchart of the steps employed for preprocessing raw hyperspectral data to evaluate snap bean yield components.

2.4.1. Calibration to Reflectance

Calibration to reflectance was executed using the three-point (white, gray, and black panels) and four-point (white, light gray, dark gray, and black) empirical line method (ELM) via ENVI software (V. 5.4) for 2019 and 2020 data, respectively. This method has proven accurate for calibrating to reflectance [57]. The reference spectra were collected using a portable spectroradiometer (“HR-1024i” model, Spectra Vista Corporation, Poughkeepsie, NY, USA).

2.4.2. Vegetation Detection

The vegetation detection step was approached via two different techniques: vegetation index (VI) and spectral angle mapper (SAM). For the VI approach, the Red edge Normalized Difference Vegetation Index (*RENDVI*) was calculated as below for each plot:

$$RENDVI = \frac{b_{750} - b_{705}}{b_{750} + b_{705}} \quad (1)$$

where b_q is the reflectance at wavelength q . Furthermore, the average of the *RENDVI* distribution was computed as the initial threshold (`thresh_initial`), where `thresh_initial` was fed to global thresholding algorithm [58] to compute the final threshold (`thresh_final`). From the *RENDVI* image, values above and below the calculated `thresh_final` were chosen as vegetation and background, respectively, and a mask was created. A median blur filter of size 7×7 was used to remove any unwanted noise from the generated mask. Finally, the generated mask was applied to all plots.

For the SAM approach, an average spectrum of 1104 collected spectra of the Huntington cultivar, from the previously published yield modelling greenhouse study [4], was used as the reference spectra [59]. Subsequently, the cosine similarity angle between each pixel and the reference spectra was calculated as:

$$\alpha = \cos^{-1} \left(\frac{\vec{s} \cdot \vec{x}}{\|\vec{s}\| \cdot \|\vec{x}\|} \right) \quad (2)$$

where s is the reference spectrum, x is the spectrum from the hyperspectral drone data, and α is the cosine similarity angle (radians). A threshold of 0.3 radians for α was iteratively selected to preserve the pixels closest to the reference spectrum. The generated mask from this approach was applied to all plots, after which the extracted vegetation spectra were fed to the spectral denoising step.

2.4.3. Spectral Denoising

The flowchart presented in Figure 4, shows that the spectral denoising stage begins with removing spectral regions between 400–450 nm and 900–1000 nm, at the detector fall off domains (i.e., two ends of the spectral collection domain), both of which exhibit a lower signal-to-noise ratio (SNR) compared to other regions. SNR was calculated by dividing the mean by the standard deviation of each spectrum. This step reduced the number of bands from 272 to 202. Next, the proposed spectral denoising algorithm from Chen et al. (2010), was implemented. First and foremost, the data from all the plots and cultivars for each day were concatenated and passed to a PCA. This data fusion step allows us to capture a holistic variability that might not be obtainable if approached by another method. Subsequently, the MAP method, introduced in the introduction section, was used to obtain the first k critical PC's. The MAP method begins with partialling out the component loading matrix A_k from the correlation matrix R , in order to obtain partial covariance matrix C_p , where k is the number of components (from 1 to p ; p = the number of variables) as below:

$$C_p = R - A_p A_p^T \quad (3)$$

the average square of the partial correlations is calculated as a measure for *MAP* criterion:

$$MAP_k = \sum_{i=1}^p \sum_{\substack{j=1 \\ i \neq j}}^p \frac{r_{ijk}^2}{p(p-1)} \quad (4)$$

where *MAP* is an array of length p . From this method, k , the number of components to retain, is determined by finding the minimum *MAP* and its corresponding k th component, i.e., $k = \text{argmin}(MAP)$ [60]. Next, $p - k$ components for each spectrum from each plot are passed to a dual-tree complex wavelet transform, in an iterative fashion. The number of

levels to perform wavelet decomposition was set at five. The thresholding formula, as presented in Chen et al. (2010) is as below:

$$d_{j,k} = d_{j,k} \max\left(0, \left(1 - \frac{thr^2}{S_{j,k}^2}\right)\right) \quad (5)$$

where d_j is the high-pass wavelet coefficient i , with size k . $S_{i,j}^2 = (|d_{j,k-1}|^2 + |d_{j,k}|^2 + |d_{j,k+1}|^2)/3$, and $thr = \sqrt{2\sigma_n^2 \log n}$ is the universal threshold. σ is the noise estimate calculated as $\sigma = MAD(|d_{1,k}|)/0.6745$, with MAD being the median absolute deviation, and $d_{0,k}$ being the high-pass coefficient at the finest level. Next, the updated coefficients are transformed back into the PCA space, concatenated with the retained PC's, and then transformed back to the spectral domain.

2.5. Data Analysis

2.5.1. Jostar: Feature Selection Library in Python

We developed Jostar, a feature selection library for Python for classification and regression tasks (see Section 2.6) for this study. This library is comprised of: (i) five meta-heuristic algorithms, namely ACO [48], GA [47], Differential Evolution (DE; [61]), and Simulated Annealing (SA; [62]); (ii) three sequential algorithms which are Plus-L Minus-R (LRS), Sequential Forward Search (SFS), and Sequential Backward Search (SBS) [63]; and (iii) one multi-objective optimization algorithm that is Non-dominated Sorting Genetic Algorithm-II (NSGAI; [64]). Among meta-heuristic algorithms, search spaces for GA, DE, and NSGAI were initially defined from an arbitrary continuous domain of uniform distribution, i.e., $x \sim U(0, 1)$. The continuous domain is then converted to discrete domain (permuted indices; suited for feature selection) using random keys defined as $y = [argsort(f)_m]_{1 \leq m \leq n_f}$ with f as the coordinates in the defined search space, and y as the random keys array of size n_f . The rest of the meta-heuristic algorithms are defined in the discrete domain, thus no further alteration was needed.

These algorithms ingest multiple parameters. The n_f parameter defines how many features need to be selected from the given feature space. However, algorithm-specific parameters lack support from the literature and the vast majority of heuristics mandate hyperparameter tuning prior to final employment. Jostar takes advantage of an easily operated function for hyperparameter tuning (see Section 2.6). Moreover, as mentioned in the introduction section, removing correlated variables is necessary, which can be performed in various ways in Jostar. The user can sort features prior to decorrelating variables based on scores obtained from statistical tests, such as the F-test, mutual information, χ -squared, and Pearson's correlation coefficient, between dependent and independent variables. Furthermore, the user has access to numerous scoring functions for minimization or maximization, such as coefficient of determination (R^2), adjusted coefficient of determination (R_{adj}^2), root mean square error (RMSE), etc. for regression tasks, and precision, recall, kappa, among others, for classification tasks.

To compare selected features from one optimization model with another, a robust ranking method needs to be implemented. Such a method not only accounts for the occurrence rate of important features during optimization, but also how much each of those features contributed to overall performance. Thus, a ranking attribute defined, using the proposed method, is available to all heuristic algorithms:

$$\begin{cases} t_m = u_m(1 + [1 - \alpha_m]w_m); & \text{if regression task} \\ t_m = u_m(1 + \alpha_m w_m); & \text{if classification task} \end{cases} \quad (6)$$

where variables are: m is 1 to p with p being the number of features passed to the optimization algorithm; t_m is the calculated feature ranking; u_m is the proposed feature permutation method [65] calculated for all features; and w_m is the occurrence rate. The occurrence rate is defined as $w_m = n_{m,occ} / n_{total}$ for feature m , for which $n_{m,occ}$ is the number of times feature m appeared in the top 25% of the features subset pool and n_{total} is the total number of feature

subsets in the top 25% percentile of the features subset pool (e.g., of 1000 agents generated during optimization, the top 25% percentile are extracted counting to 200 agents [$n_{total} = 200$], and the number of times feature m occurred in n_{total} set is logged as $n_{m,occ} = 150$; thus the w_m is calculated as $w_m = n_{m,occ} / n_{total} = 150 / 200 = 0.75$); and α_m is the average $RMSE$ score for regression tasks and mean F_1 -score for classification tasks, for feature m . This last step takes into account the importance of number of times a feature appears in the search domain, for feature subsets of low error or high accuracy (i.e., the top 25%). Moreover, the α_m parameter, multiplied by w_m , takes into account the impact the feature m has when compared to the perfect accuracy, i.e., $RMSE =$ or F_1 -score = 1, therefore enabling a ranking comparison from different optimization models for the same data set.

2.5.2. Feature Selection Procedure

The overall data analysis procedure is depicted as a flowchart in Figure 5. As can be seen, the preprocessed hyperspectral data are initially fed to the feature selection step in hyperparameter tuning (param. tuning) mode, with `max_evals = 200` as the maximum number of evaluations, to search the defined space and `n_iter = 20` iterations for the optimization model, at each evaluation. Subsequently, the test mode is run with the optimal parameter identified from the param. tuning mode, with `n_iter = 100`. For both modes, the following steps were conducted: (i) PLSR was selected as the regressor for this study's objective, given its capability for handling multicollinearity and its low computational cost; (ii) GA, SA, PSO, and ACO, embedded in Jostar, were selected as four feature selection optimization models; (iii) the defined number of selected features was set as five (i.e., $n_f = 5$) for all sets and cross-validation Leave-One-Out $RMSE$ value was selected as the scoring parameter and minimized for (`weight = -1`); (iv) the top 20 decorrelated features were retained, in order to ensure a consistent search space across all tests; (v) all features were scaled to the 0–1 range; (vi) the optimization model was then “fit” with the defined arguments and values. The test mode takes advantage of averaging output rankings from the four optimization models. The last step allows for a more holistic ranking, since there is a possibility of an optimization algorithm falling in a local minimum. The defined search spaces for the hyperparameter tuning step are tabulated in Table 2. More information on each parameter's explanation can be found in Jostar's documentation.

Table 2. Optimization models, corresponding parameters, and associated value ranges identified for hyperparameter tuning.

Optimization Model	Parameter	Description	Sampling Method	Low Bound	High Bound
GA ¹	$Cross_{perc}$	Crossover percentage	Uniform logarithmic	10^{-2}	0.5
	Mut_{perc}	Mutation percentage	Uniform logarithmic	10^{-2}	0.2
	Mut_{rate}	Mutation rate	Uniform logarithmic	10^{-2}	0.2
	β	Selection pressure	Random integer	1	10
	N_{pop}	Population size	Random integer	20	200
SA ²	α	Cooling factor	Uniform logarithmic	0.8	0.99
	T_0	Initial temperature	Uniform logarithmic	10^{-1}	500
	$N_{iter-sub}$	Number of sub-iterations	Random Integer	20	200
PSO ³	α	Information elicitation factor	Uniform logarithmic	10^{-5}	0.5
	C_1	Cognitive parameter	Uniform logarithmic	10^{-2}	2
	C_2	Social parameter	Uniform logarithmic	10^{-2}	2
	W	Inertia weight	Uniform logarithmic	10^{-2}	1.2
	W_{damp}	Inertia weight damping factor	Uniform logarithmic	10^{-2}	0.5
	N_{pop}	Number of particles	Random integer	20	200
ACO ⁴	α	Information elicitation factor	Uniform logarithmic	10^{-5}	0.5
	ρ	Pheromone evaporation coefficient	Uniform logarithmic	10^{-5}	0.5
	τ_0	Initial pheromone intensity	Uniform logarithmic	10^{-5}	1
	Q	Pheromone intensity	Uniform logarithmic	10^{-5}	1
	N_{ant}	Number of ants	Random integer	20	200
	β	Meta-heuristic factor	Random integer	1	5

¹ Genetic Algorithm; ² Simulated Annealing; ³ Particle Swarm Optimization; ⁴ Ant Colony Optimization.

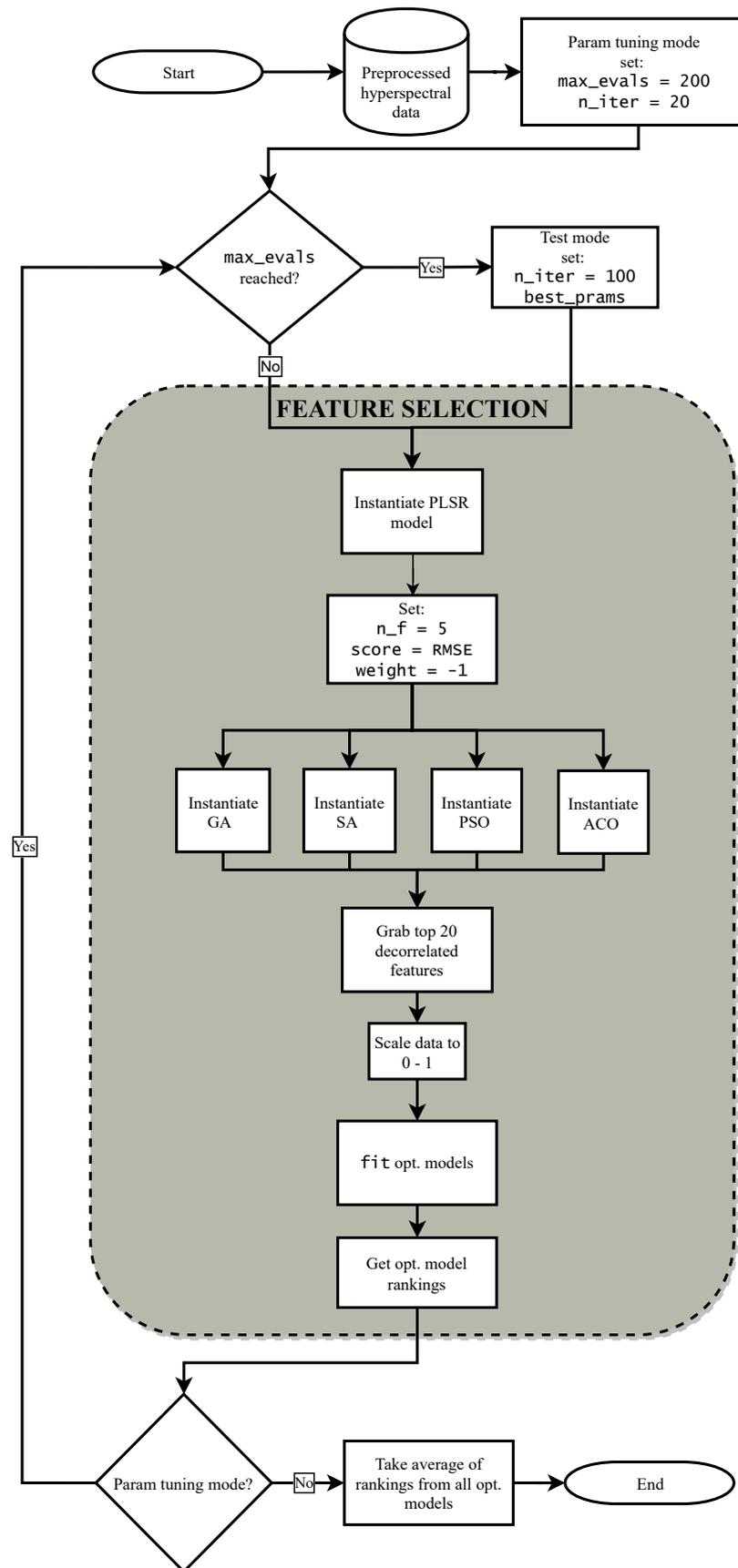


Figure 5. Flowchart of the steps employed for data analysis of the preprocessed hyperspectral data.

2.6. Software

Jostar was implemented and tested in Python 3.x [66] and is publicly available at www.github.com/yxoos/jostar. The hyperparameter tuning function was implemented with the help of Hyperopt [67], while the PLSR was sourced from Scikit-learn [68].

3. Results

3.1. Descriptive Statistics

Box plot representations of ground truth data for both pod weight and seed length yield indicators are depicted in Figure 6. Pod weight had a larger range, average and similar or larger variability in 2019 compared to 2020 (Figure 6a,b). Seed length range, average and variability were higher for data collected in 2020 (Figure 6c,d).

Figure 7 depicts the SNR comparison between the raw data prior to denoising, and the denoised data (PCA denoised + band removal), from a random flight from the 2020 data set. As can be seen, there was a significant increase in the reported average SNR (21% increase) post denoising, especially in the visible region. Moreover, the plot shows that SNR dropped in detector fall-off regions—400–450 nm and 900–1000 nm—further justifying our approach to remove the bands from the two extreme ends of collected spectra.

The impact plots (see [69]) of the first 10 PC's are shown in Figure 8. PC-1 emphasized changes in the NIR region and the variability in terms of amplitude, while PC-2 highlighted variability in VIS and NIR, but more in terms of a *slope*, with the handle set at red edge peak region (i.e., higher/lower shifts on one end and lower/higher shifts on the other end). The first two PC's accounted for 65% of total explained variability. PC-3 and PC-4 were responsible for a total of 8% explained variability and mostly reflected small variabilities in the visible region. PC-5 was responsible for only slight changes in the VIS domain and accounted for 1.32% of the total variability. The higher the PC index, the smaller the variability described. PC-6 to PC-8 reflected noise in the VIS and the far end of the NIR region.

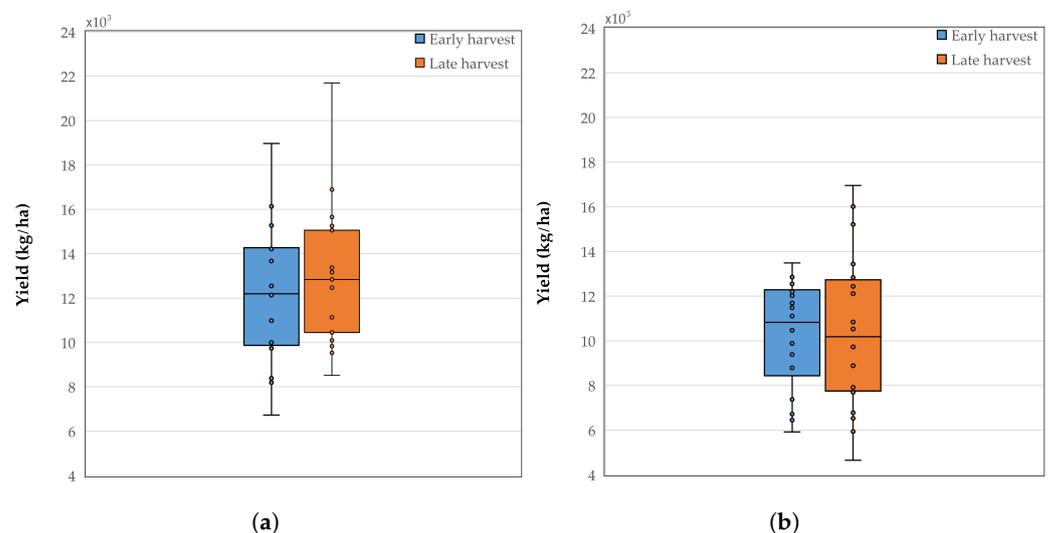


Figure 6. Cont.

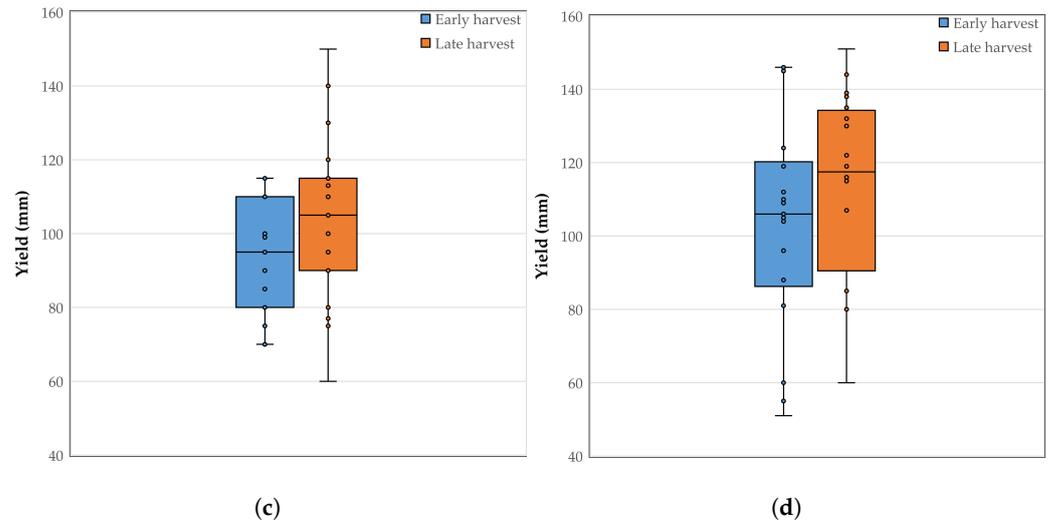


Figure 6. Ground truth yield values for pod weight and seed length for 2019 and 2020 data: (a) pod weight 2019 data, (b) pod weight 2020 data, (c) seed length 2019 data, and (d) seed length 2020 data.

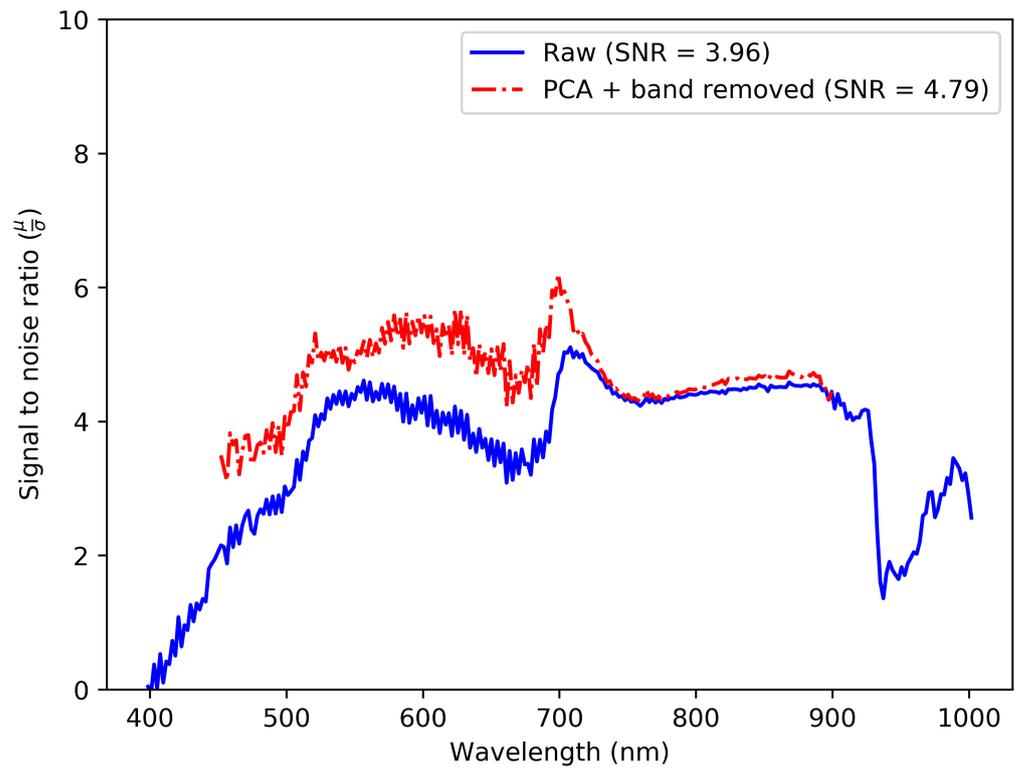


Figure 7. A comparison of signal-to-noise (SNR) across wavelengths before and after spectral denoising. Note the increase in the average estimated SNR.

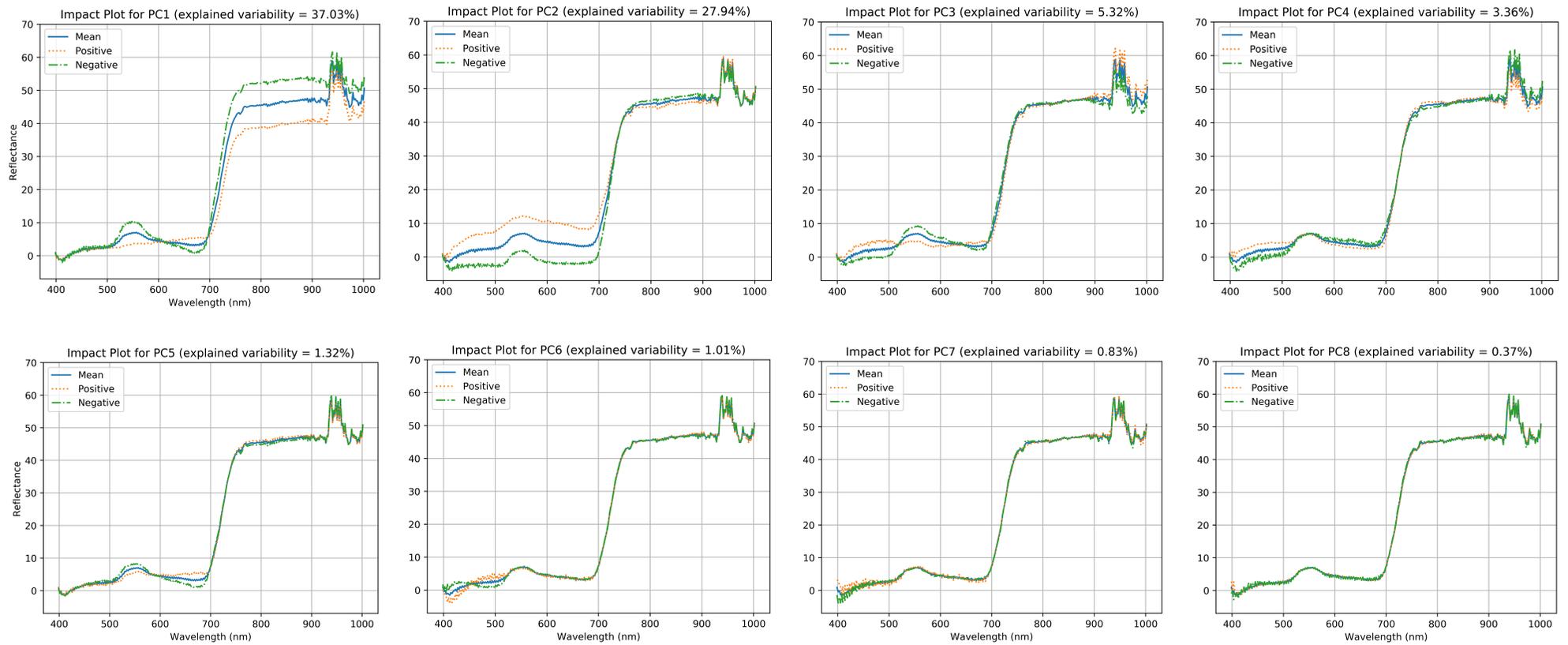


Figure 8. The first 10 principal components of a random flight from the 2019 data set. Please note that the first few PC's contribute to the largest amount explained variability.

3.2. Pod Weight

3.2.1. 2019 Data Set

Among optimization models, ACO and SA outperformed others at both early and late harvest stages (Figure 9). PSO became stuck at a local minimum for both harvest stages. We observed a maximum R^2 of 0.78, reported by the ACO-RENDVI at 46 DAP for early harvest, and a maximum R^2 of 0.85 at 48 DAP, reported by SA-SAM, for late harvest. Moreover, it seemed that the RENDVI model had a larger variability compared to SAM. The general trend of both curves for all models, however, was of an ascending nature. We noted that for both vegetation models, there was a drop in accuracy at 39 DAP, which corresponded to the final stages of flowering. Figure 10 displays regression results for the ACO-RENDVI model at 46 DAP and SA-SAM model at 48 DAP, for early and late harvest, respectively. The early harvest 2019 regression results exhibited an R^2 as high as 0.78 for calibration and an R^2 of 0.60 for cross-validation. The corresponding RMSE values were 1369 kg/ha (12.9%) and 1838 kg/ha (17.2%) for calibration and cross-validation, respectively. The late harvest stage for the 2019 data set outperformed the early harvest stage with an R^2 of 0.84 and 0.75, for calibration and cross-validation, respectively. The corresponding RMSE values for the late harvest stage of the 2019 data set were 1257 kg/ha (10.7%) and 1588 kg/ha (13.8%). Residual plots from both harvest stages demonstrated a random spread without a trend, validating the choice of the regression model.

We chose the SAM models at both harvest stages for further analysis toward discriminating bands, due to the lower variability in terms of performance metrics throughout the growth stage. Figure 11 shows the normalized average rankings across all optimization models and the corresponding top five features to the right of each ranking "belt" for the 2019 pod weight data. Figure 11a shows rankings for the early harvest stage. Blue reflective pigment (450–500 nm) was present on all days, except at 50 DAP; the green reflective pigment (500–560 nm) was represented throughout; the red reflective pigment (600–700 nm) manifested throughout, except at 39 and 46 DAP; and the red edge peak (700–770 nm) was evident throughout, except at 48 and 54 DAP. For this harvest stage, the wavelength at 451 nm was repeated three times, and 478, 505, 516, 525 nm were repeated twice in the early harvest stage analysis. Figure 11b depicts normalized averaged rankings for late 2019 harvest stage. The identified bands from these set of rankings were: blue reflective pigment (450–500 nm) was present at all dates except 35 and 50 DAP, corresponding to late flowering and pod formation; green reflective pigment (500–600 nm) was evident at all days; red reflective pigment (600–700 nm) was apparent at all stages but 39 DAP, which corresponds to final stages of flowering; the red edge peak (700–770 nm) was evident for all days, except at 46, 54 and 56 DAP; and NIR (819 nm) resided at 35 and 56 DAP, was represented, with the timing corresponded with the on-set of flowering and final stages of pod formation. At this harvest stage, the wavelength at 525 nm was repeated four times, 451 nm was repeated three times, and 505, 541, 659, 759, and 819 nm were repeated twice.

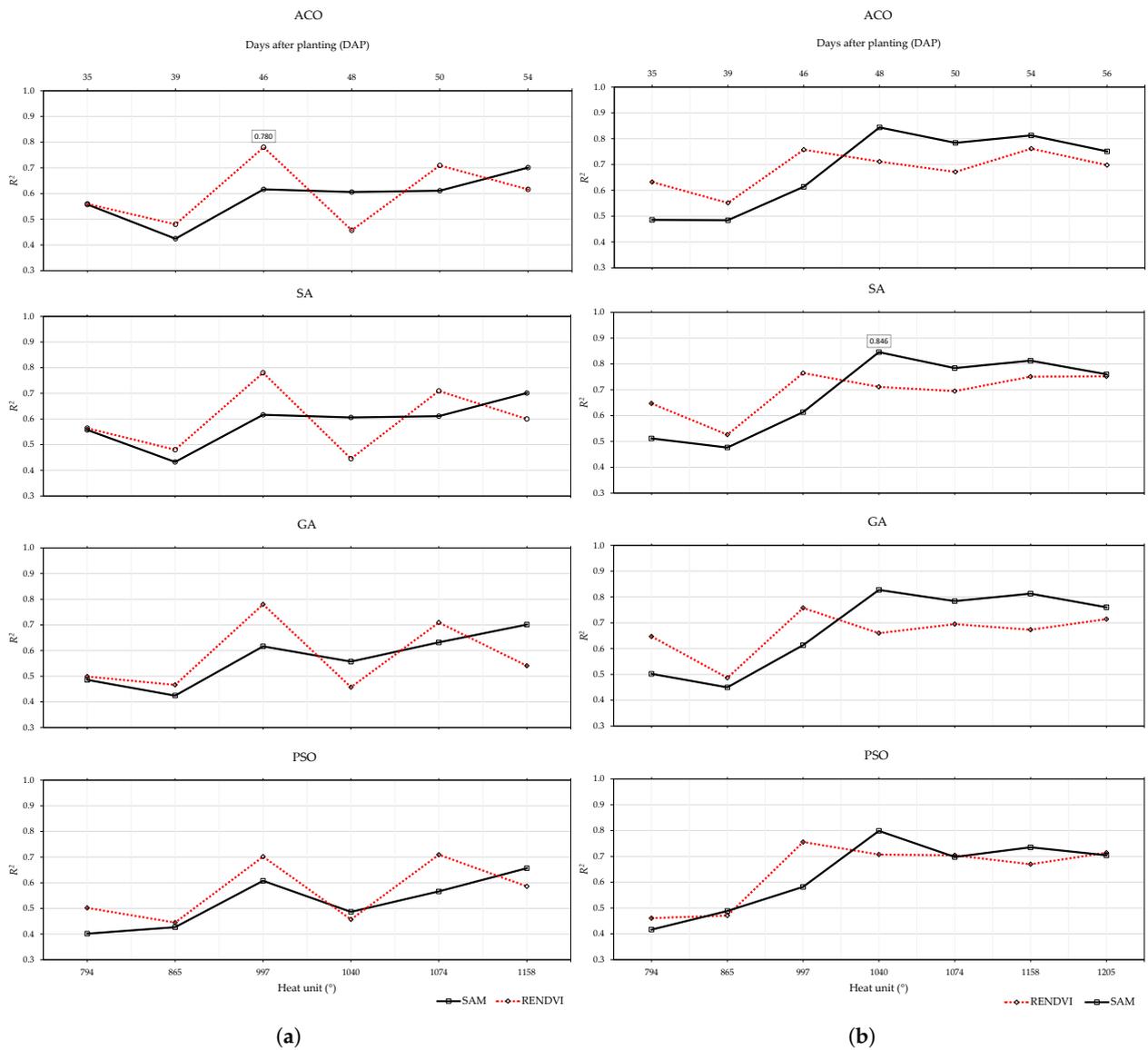


Figure 9. Results for coefficient of determination (R^2) for snap bean 2019 pod weight data set, for (a) early and (b) late harvest stage, as determined via two different vegetation approaches, SAM and RENDVI, as tested on PLSR.

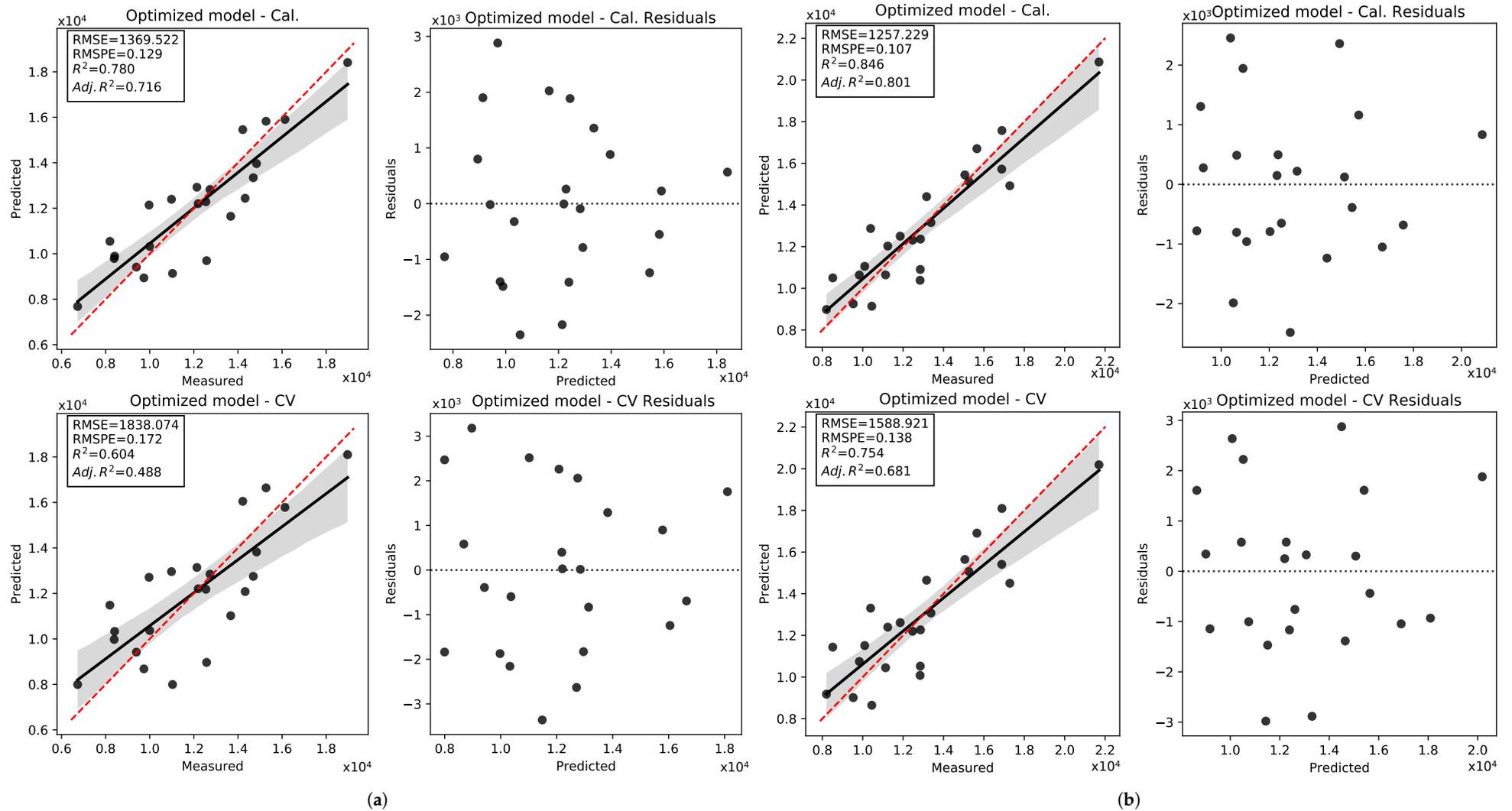


Figure 10. Regression results for snap bean 2019 data set on (a) the ACO-RENDVI model at 46 DAP for the early harvest stage and (b) the SA-SAM model at 48 DAP for the late harvest stage.

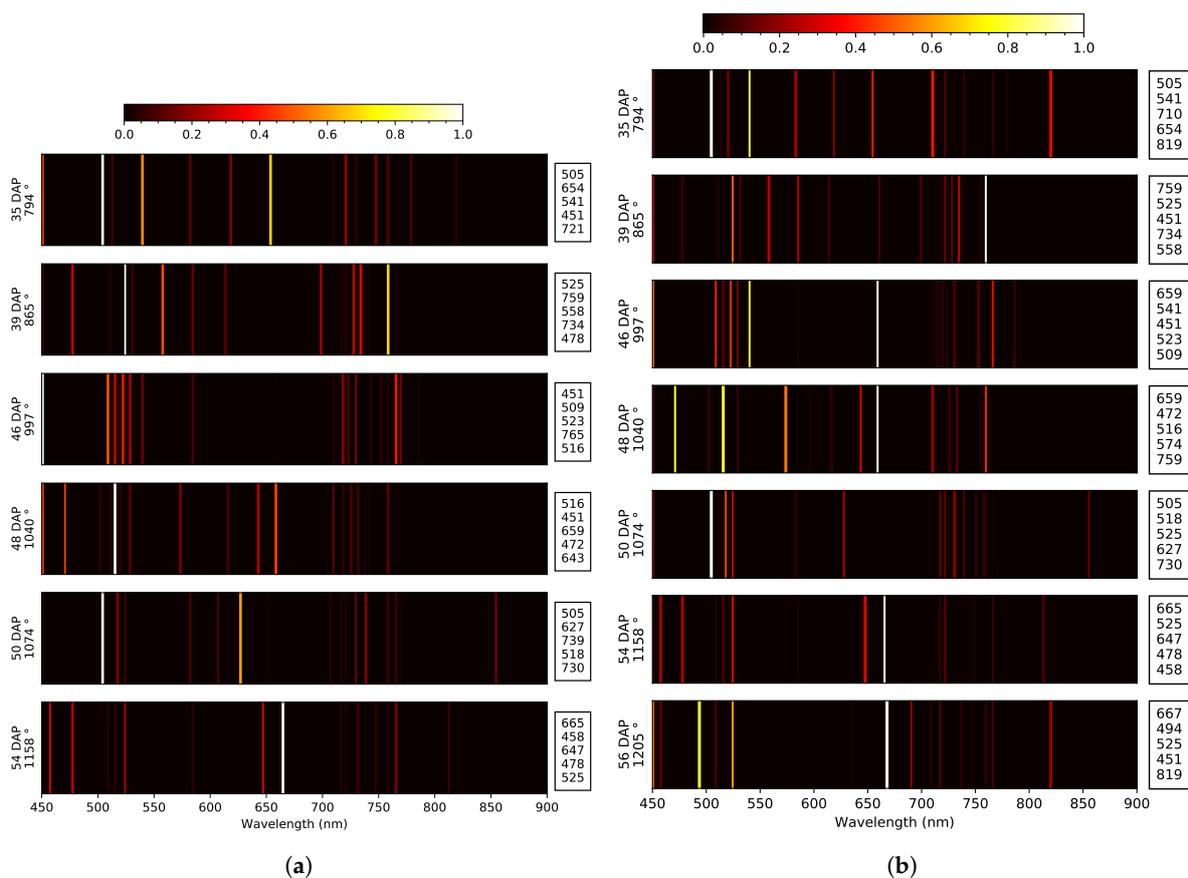


Figure 11. Snap bean pod weight 2019 data set average rankings at (a) the early harvest stage and (b) the late harvest stage, based on the SAM vegetation model. The corresponding days after planting (DAP) and heat unit for the collected data are shown to the left of each ranking belt, while the top five features, with the first one having the highest score, are listed to the right.

3.2.2. 2020 Data Set

Results for this data set are presented in Figure 12. The overall trend of plots is ascending for both harvest stages, although, a drop in accuracy for the final stage of flowering at 34 DAP was observable for both harvest stages. Moreover, it should be noted that there was a second drop in accuracy at 55 DAP for early harvest, which was not evident in the late harvest stage. ACO outperformed other optimization models, however, at 40 DAP, ACO performed inferior to other optimization models. This corroborated our reasoning for using the proposed ranking method, i.e., to account for variability in performance between different optimization models due to becoming stuck in a local minimum. SAM produced more accurate results, compared to the RENDVI model in most cases, for both early and late harvest stages. For the early harvest stage, ACO-SAM at 44 DAP outperformed other models, and for the late harvest stage, ACO-SAM at 58 DAP exhibited the most accurate results (Figure 13). Pod weight prediction at the early harvest stage was quantified with an R^2 of 0.82 and an $RMSE$ value of 987 kg/ha (11%), for the calibration data set, and an R^2 of 0.59, with corresponding $RMSE$ value 1492 kg/ha (17.5%) for cross-validation models. There was an increase in the overall performance for the late harvest stage (Figure 13b), with an R^2 of 0.93 and an $RMSE$ of 940 kg/ha (14.1%) for calibration, and R^2 of 0.84 with a matching $RMSE$ of 1377 kg/ha (19.4%) for cross-validation. The distribution of residuals was random (Figure 13b).

For the 2020 pod weight data, the SAM vegetation detection model was selected due to more consistent results with less variability, when compared to the RENDVI vegetation approach (Figure 14). The wavelengths at the early harvest stage were identified as: the blue reflective pigment (450–480 nm) was present throughout; the green reflective pigment (510–600 nm) presented throughout the experiment, except for 44 DAP; the red reflective

pigment (600–700 nm) was evident for all days except, at the harvest date (58 DAP); and the red edge peak was present throughout (700–760 nm). At this harvest stage, wavelengths at 451 and 759 nm were repeated five times each, 523 nm was repeated three times, and 518, 607, 647 nm were repeated twice at the early harvest stage. Highly ranked wavelengths detected at the late harvest were: those related to the blue reflective pigment (450–490 nm) throughout, except for 58 DAP; the green reflective pigment (500–560 nm), evident on all days except for 44 DAP; the red reflective pigment (600–690 nm), observable throughout except for 40 DAP (initial pod formation stage); and the red-edge peak (700–765 nm) observable throughout except for 31, 55, and 60 DAP. For this harvest stage, bandwidths at 451 and 759 nm were repeated four times, and 474, 500, 525, 654, and 707 nm were all repeated twice at the late harvest stage.

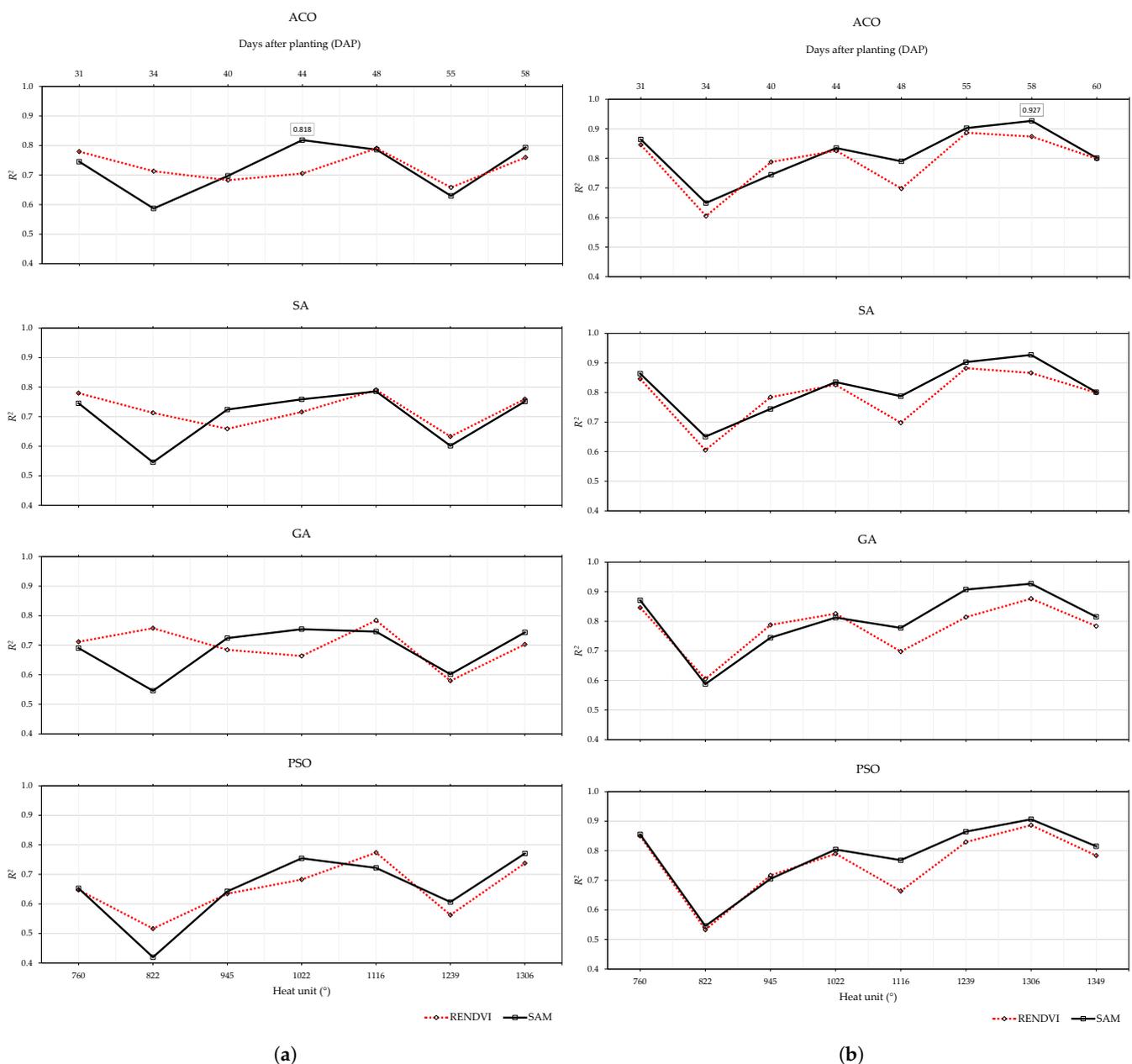


Figure 12. Coefficient of determination (R^2) results for snap bean 2020 pod weight data set, for (a) early and (b) the late harvest stage via PLSR model and two different vegetation approaches, SAM and RENDVI.

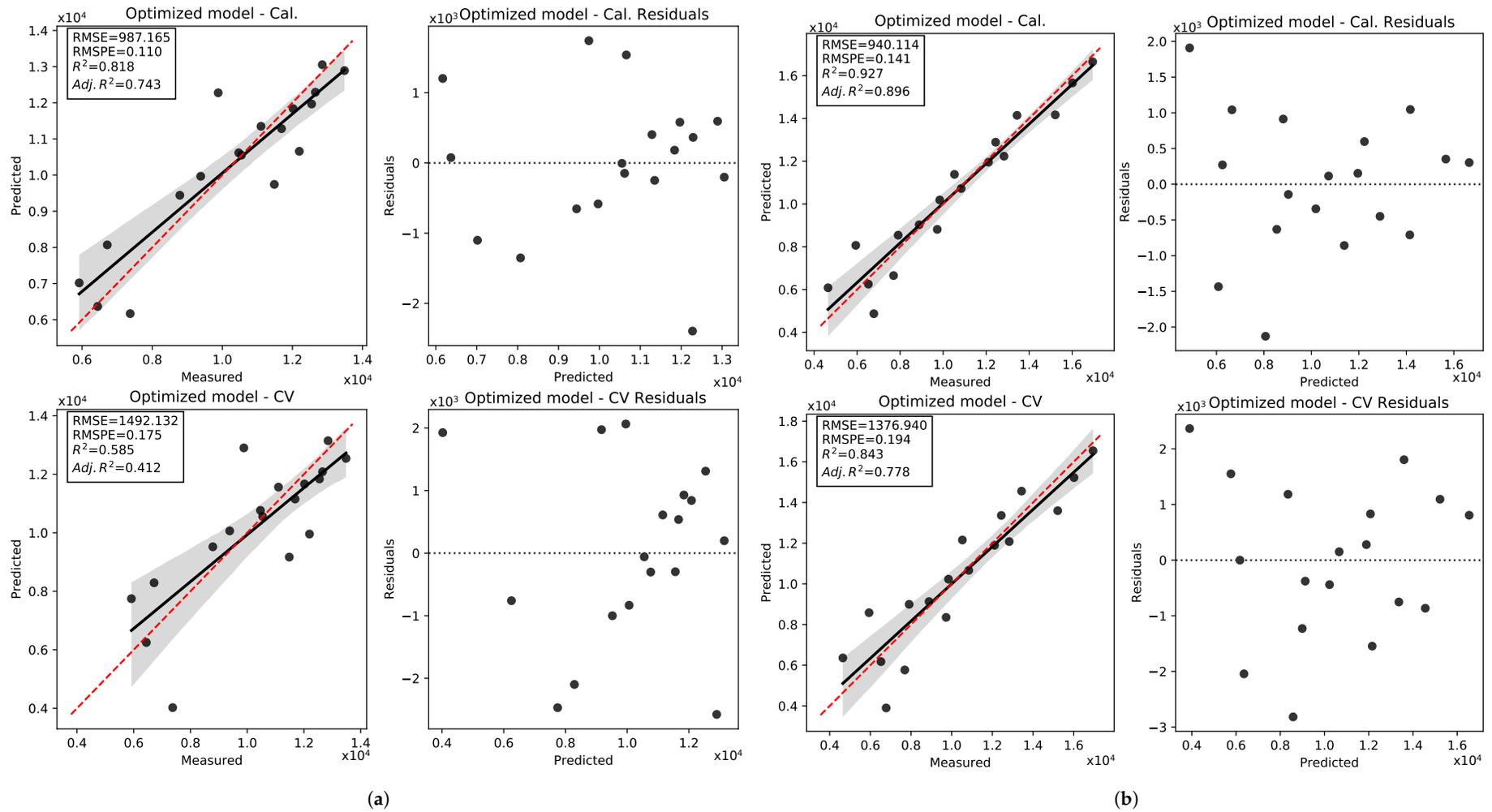


Figure 13. Regression results for snap bean 2020 data set on (a) the ACO-SAM model at 44 DAP for early harvest stage and (b) the ACO-SAM model at 58 DAP for the late harvest stage.

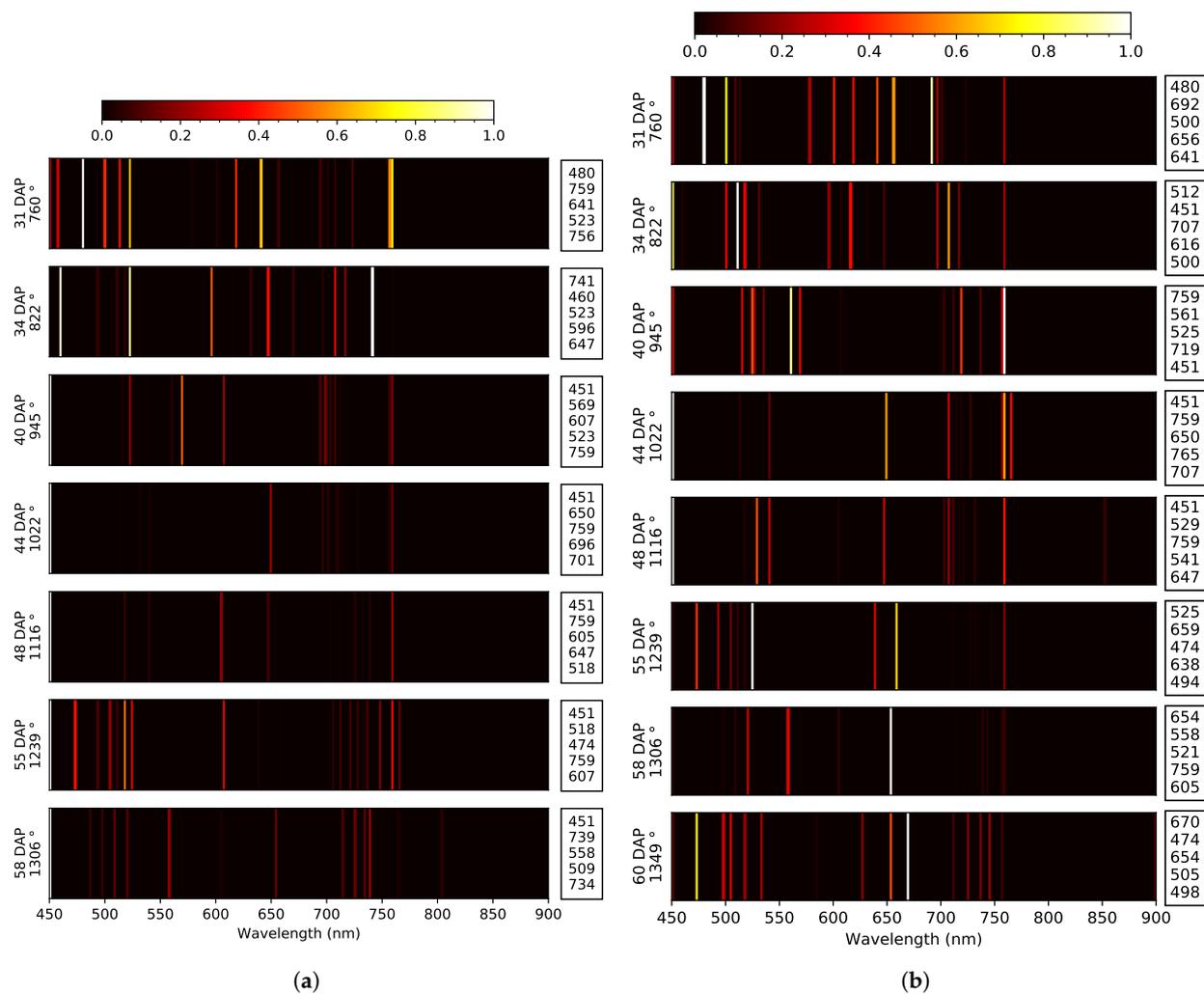


Figure 14. Snap bean pod weight the 2020 data set average rankings across all four optimization models for SAM model at (a) the early harvest stage and (b) the late harvest stage. The corresponding days after planting (DAP) and heat unit for the collected data are shown to the left of each ranking belt, while the top five features, with the first one having the highest score, are listed to the right.

3.3. Seed Length

3.3.1. 2019 Data Set

The 2019 snap bean seed length model slightly outperformed the pod weight 2019 data set (Figure 15). Similar to pod weight, the 2019 seed length results also exhibited a drop in accuracy at the initial stages of data collection, e.g., 39 DAP for early and 46 DAP for the late harvest stage. SAM outperformed the RENDVI approach by a small margin and was more consistent than RENDVI. For both early and late harvest stages, we observed a decrease in accuracy on harvest date; however, the general trend was ascending. A comparison of optimization models showed that, holistically, ACO outperformed other approaches. GA-RENDVI's results at 50 DAP surpassed other model performance, for the early harvest model (Figure 15a). Moreover, results from the late harvest stage (see Figure 15b) showed that SA-RENDVI resulted in the highest accuracy at 56 DAP. We chose to present regression results for the two mentioned models (i.e., GA-RENDVI at 50 DAP and SA-RENDVI at 56 DAP, for early and late harvest stages, respectively) in Figure 16. We conclude from this figure that the GA-RENDVI model performed accurately, with an R^2 of 0.83 and an RMSE of 6.018 mm (6.9%) for the calibration set, and an R^2 of 0.72 and an RMSE of 7.577 mm (8.7%) for the cross-validation set. The SA-RENDVI model for the late

harvest stage exhibited a higher R^2 of 0.84 and an $RMSE$ of 9.404 mm (9.2%) for calibration set, while the cross-validation yielded an R^2 of 0.73 and a $RMSE$ of 12.383 mm (12.2%). The residual plots for both of these models again showed a random spread and did not follow any trend.

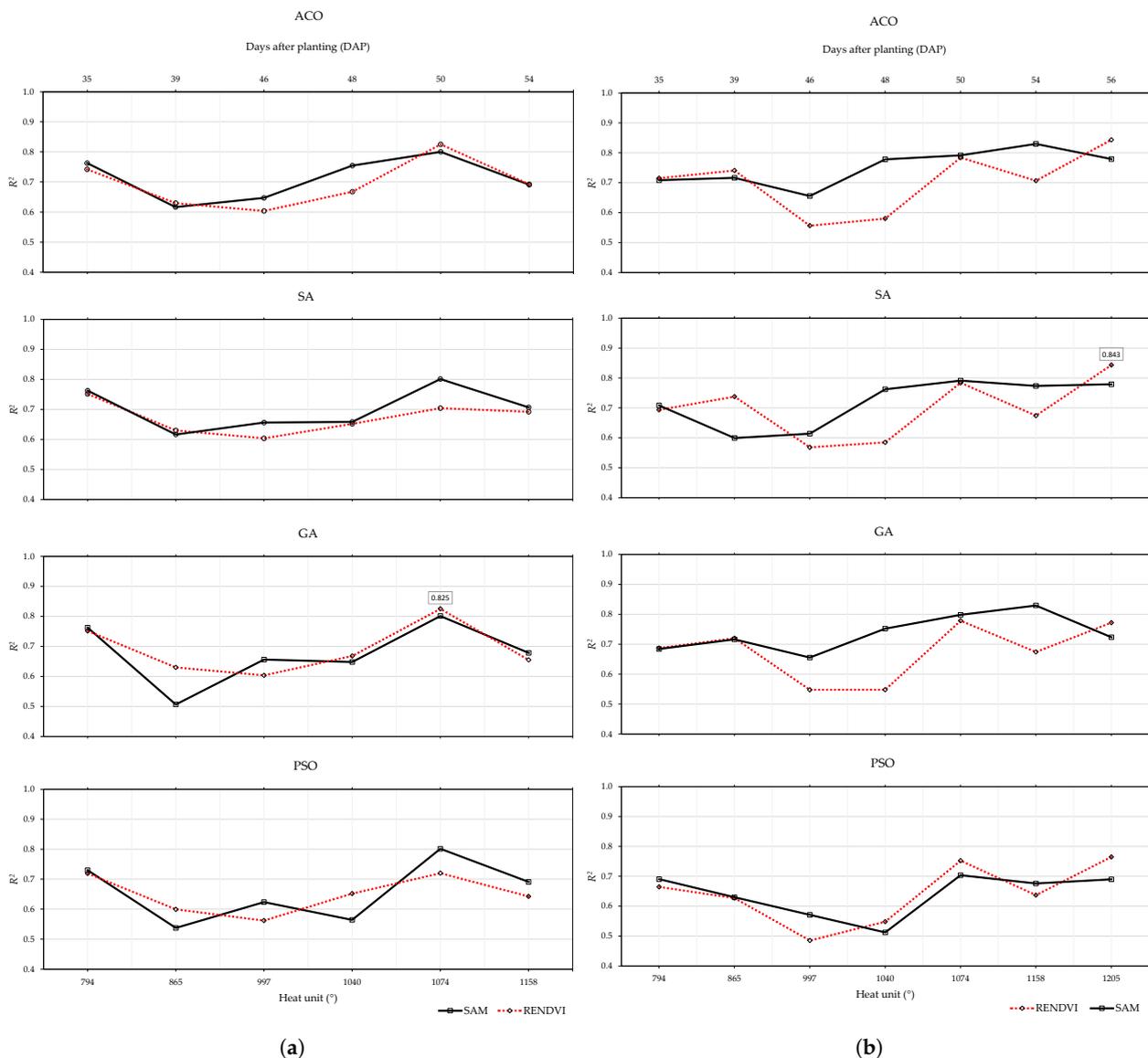


Figure 15. Results for the coefficient of determination (R^2) for snap bean 2019 seed length data set, for (a) the early and (b) the late harvest stage, via two different vegetation approaches, namely SAM and RENDVI, as tested via PLSR.

Next, we opted to show ranking results for early and late 2019 seed length data sets for the SAM vegetation approach, which resulted in lower variability and more consistency in terms of accuracy at both harvest stages (see Figure 17). We concluded the following for the early harvest stage, in terms of harvest-specific wavelengths: the blue reflective pigment (450 nm) was represented only at 39 and 48 DAP; the green reflective pigment (510–585 nm) was present throughout, except at 50 and 54 DAP (last two days); the red reflective pigment (620–700 nm) was apparent throughout; the red edge peak (700–765 nm) was evident for all stages, except at 35 and 48 DAP; and the NIR region (810–860 nm) only manifested at 50 and 54 DAP (the last two days). The wavelengths at 451, 585, 659, 710, and 716 nm were repeated twice for this harvest stage. Moreover, the bands identified for the late harvest stage model included: the blue reflective pigment (450–460 nm) for all stages,

except at 35 and 50 DAP; the green reflective pigment (510–585 nm) throughout, except for the last two days (at 54 and 56 DAP); the red reflective pigment (620–700 nm) for all except at 54 DAP; the red-edge peak (700–765 nm), except at 48 DAP; and the NIR region (810 nm), but only at 54 DAP. Further scrutiny showed that the wavelengths at 451 and 721 nm were repeated three times, and those at 583, 585, 659, 716 nm were repeated twice for the late harvest stage.

3.3.2. 2020 Data Set

Figure 18, in turn, shows results for the 2020 data set, which exhibited a dip at 44 DAP at both harvest stages. However, the overall trend was ascending for both stages. The performance metrics throughout the growth period were higher than those reported for the 2019 data set, from a holistic perspective, with observed $R^2 > 0.75$. ACO outperformed other optimization models, but the SA-SAM model yielded the most accurate results on the day of harvest for the early harvest stage. Findings for the SA-SAM model at the early harvest stage showed an R^2 of 0.98 at 58 DAP. In contrast, the ACO-SAM model outperformed other models and resulted in an $R^2 = 0.94$ at 58 DAP for the late harvest stage. The small reduction in accuracy on the harvest date, previously reported with the 2019 seed length data set, was also detected with the 2020 late harvest stage data set. Regression results for the SA-SAM at 58 DAP for the early harvest stage and the ACO-SAM at 58 DAP for the late harvest stage are presented in Figure 19. A solid performance for the early harvest stage (Figure 19a) was evident with an R^2 of 0.98 and an RMSE of 4.245 mm (4.6%) for the calibration set and an R^2 of 0.95 and an RMSE of 6.560 mm (7.4%) for the cross-validation set. Findings for the late harvest stage (see Figure 19b) performed similarly, with an R^2 of 0.94 and an RMSE of 6.335 mm (6.7%) for the calibration set and R^2 of 0.86 with an RMSE of 9.463 mm (10.3%) for the cross-validation set. There again was no trend present in residual plots and there was a random spread observable at both harvest stages.

Figure 20 shows the ranking belts for the 2020 data set for both harvest stages. Wavelengths identified at the early harvest stage were: the blue reflective pigment (450–475 nm), observable throughout; the green reflective pigment (520–560 nm) region, visible on all dates except at 58 DAP (harvest date); the red reflective pigment (600–700 nm) spectral region, observable throughout, except at 44, 55, and 58 DAP; the red edge peak (700–740 nm) region, apparent throughout, except at 31 and 48 DAP; and the NIR (803 nm) range, which was only observable at the harvest date (58 DAP). It should be noted that 451 nm was repeated six times, and the 494 and 518 nm wavelengths were repeated twice in the analysis. The detected wavelengths at the late harvest stage, however, were: the blue reflective pigment (450–500 nm) region, apparent for all dates; the green reflective pigment (500–570 nm) range was detected throughout, except at 58 DAP, where maximum accuracy was obtained (see Figure 18); the red reflective pigment (600–700 nm) spectral region, but not at 55 and 58 DAP; the red edge peak (700–760 nm) feature, observable throughout, except at 31 and 34 DAP; and the NIR spectral feature at 803 nm, but only at 58 DAP and with the lowest significance. At this harvest stage, 451 nm was repeated seven times, and 494, 518, 670, 699, and 759 nm were repeated twice.

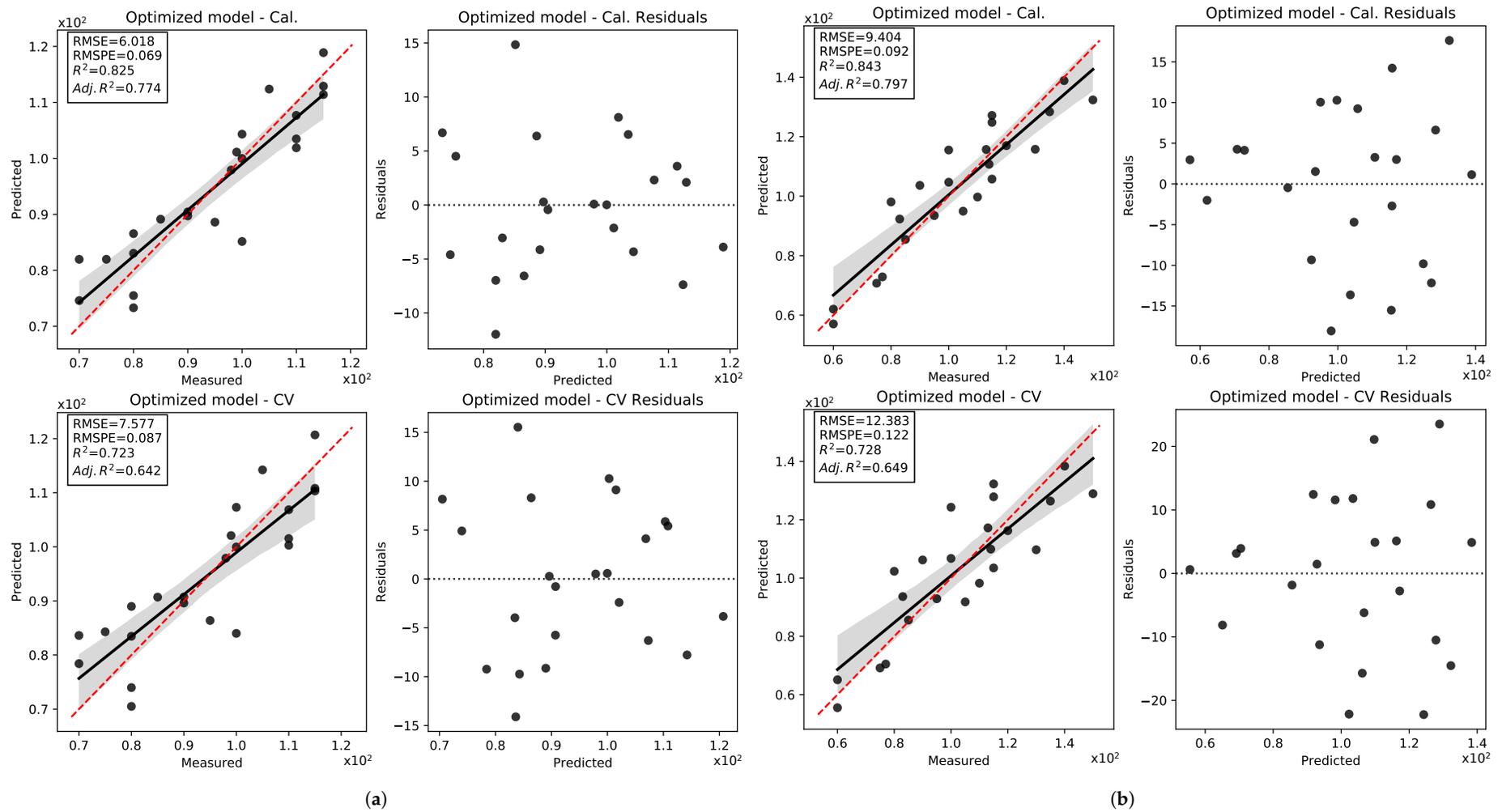


Figure 16. Regression results for snap bean 2019 seed length data set on (a) the GA-RENDVI model at 50 DAP for the early harvest stage and (b) the SA-RENDVI model at 56 DAP for the late harvest stage.

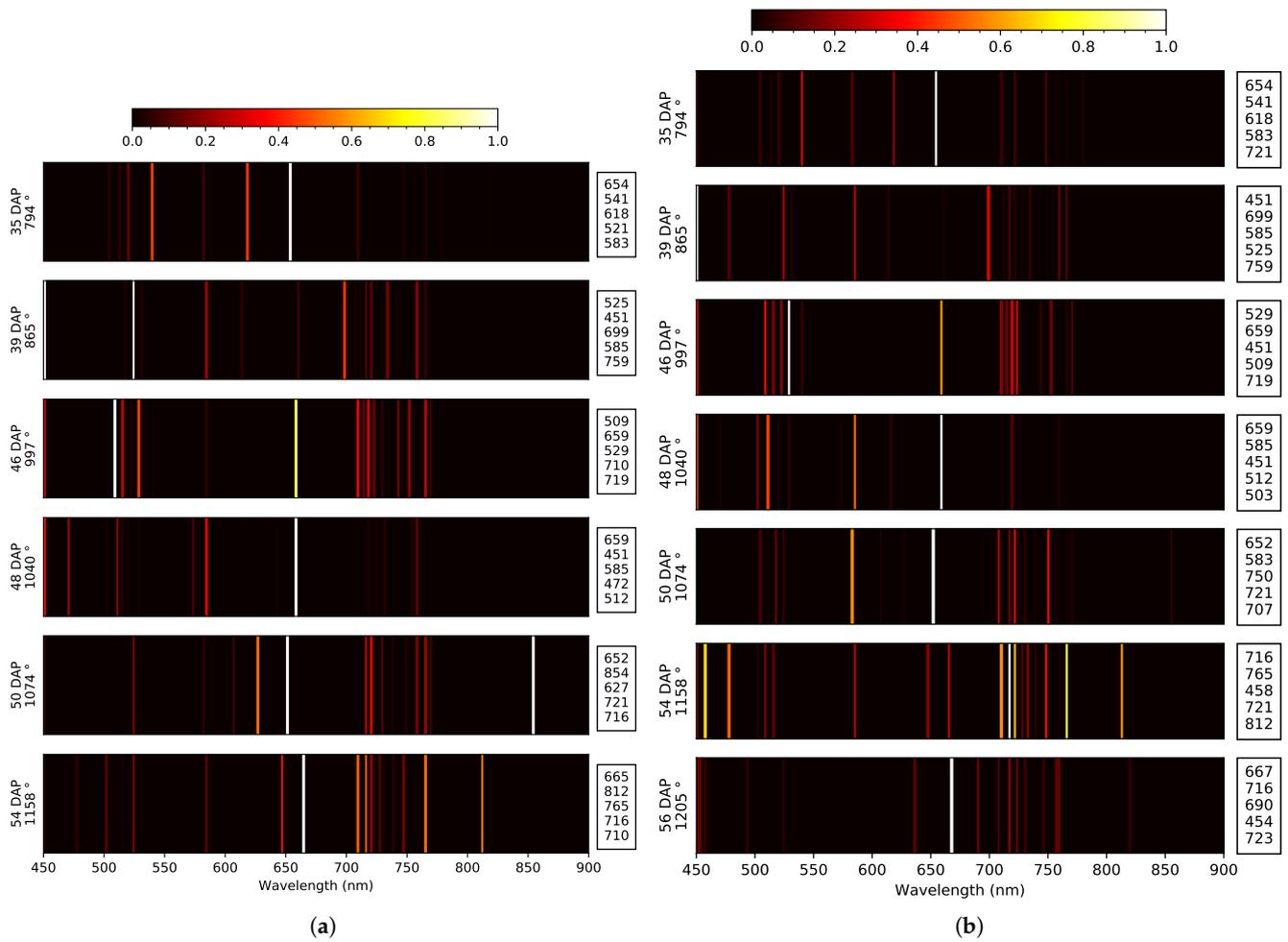


Figure 17. Snap bean seed length 2019 average rankings across all four optimization models for SAM model at (a) the early harvest stage and (b) the late harvest stage. The corresponding days after planting (DAP) and heat unit for the collected data are shown to the left of each ranking belt, while the top five features, with the first one having the highest score, are listed to the right.

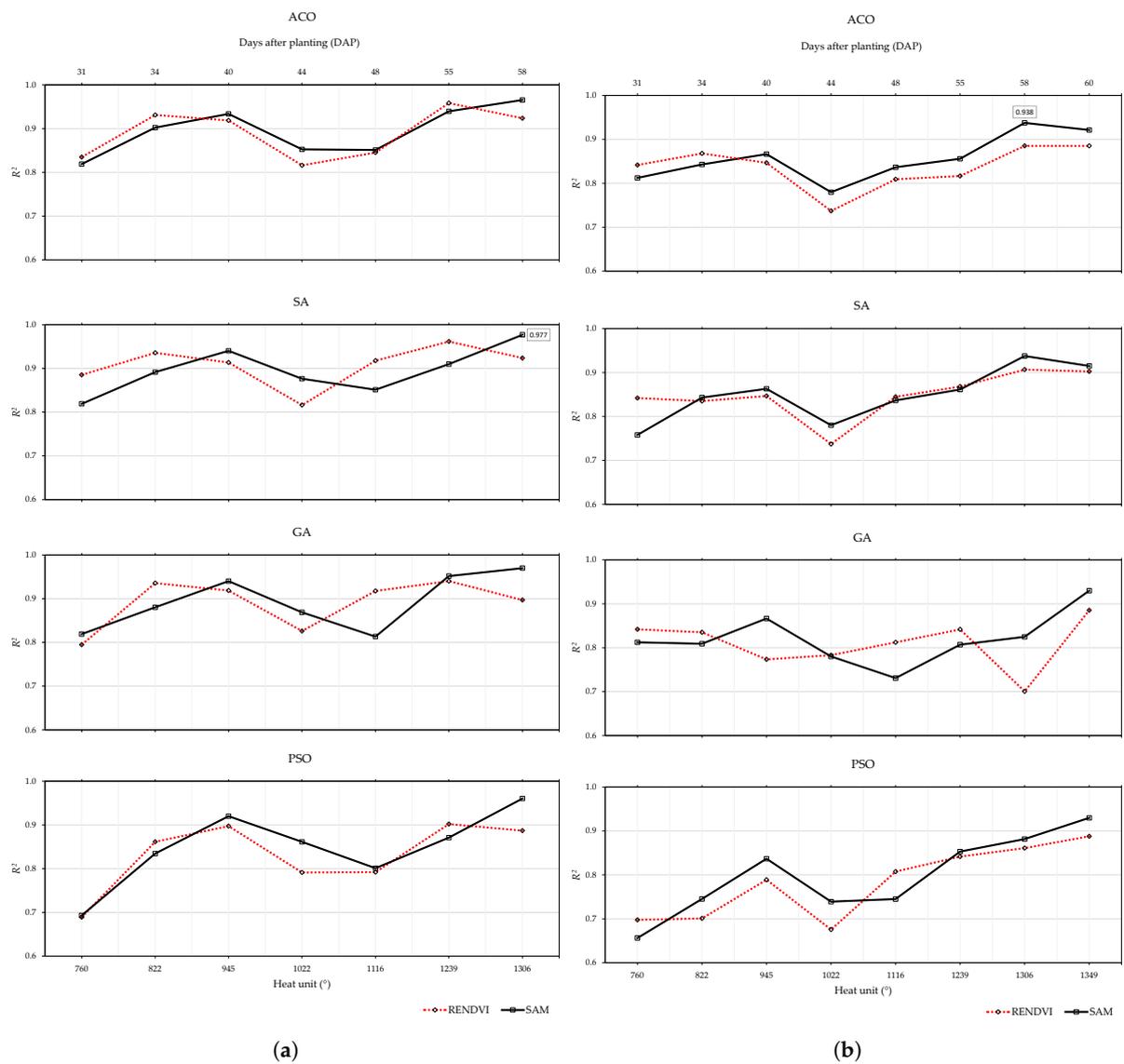


Figure 18. Coefficient of determination metric for snap bean 2020 seed length data set, for (a) early and (b) the late harvest stage via two SAM and RENDVI, tested with PLSR.

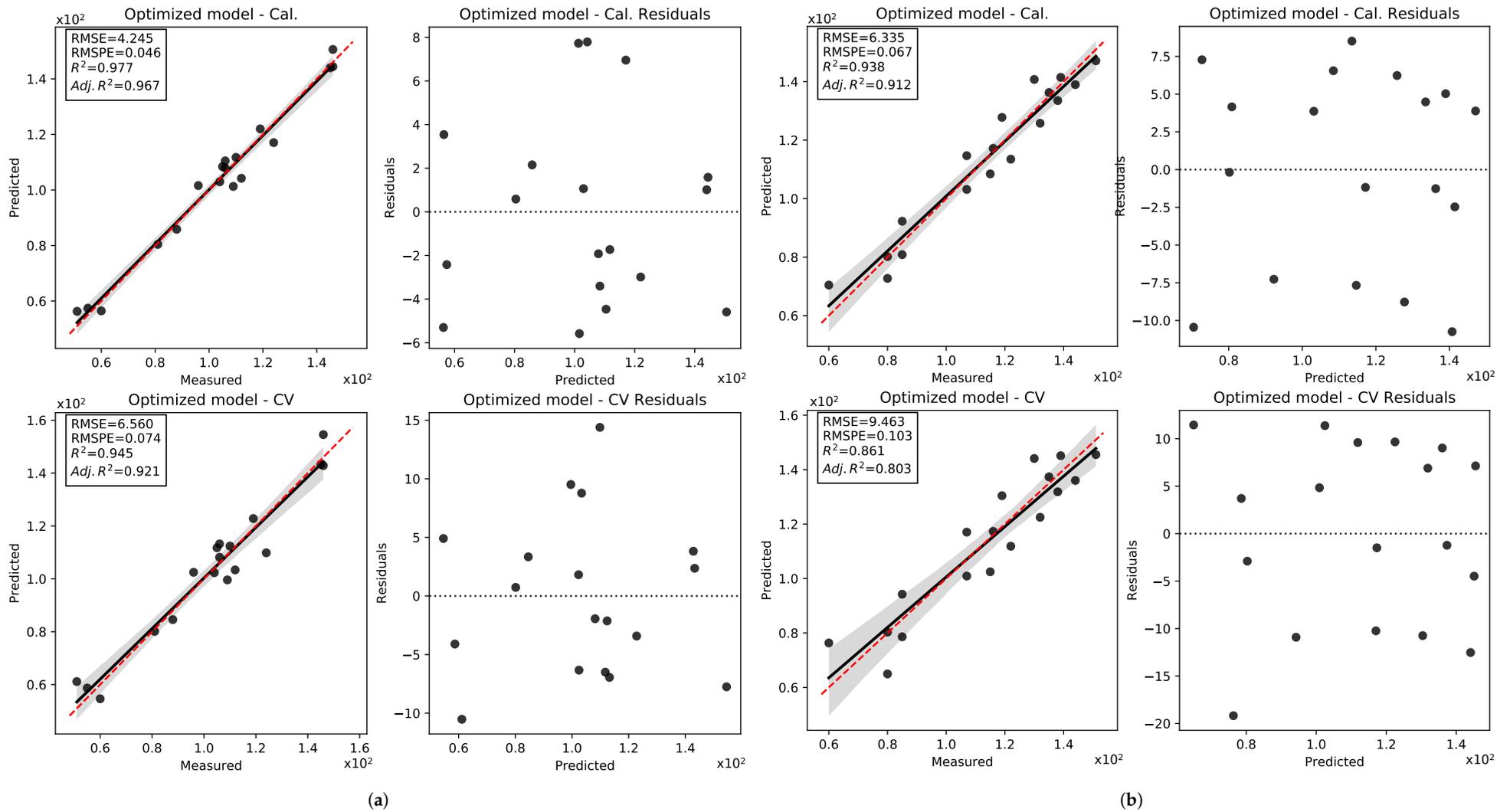


Figure 19. Regression results for snap bean 2020 seed length data set on (a) SA-SAM model at 58 DAP for the early harvest stage and (b) ACO-SAM model at 58 DAP for the late harvest stage.

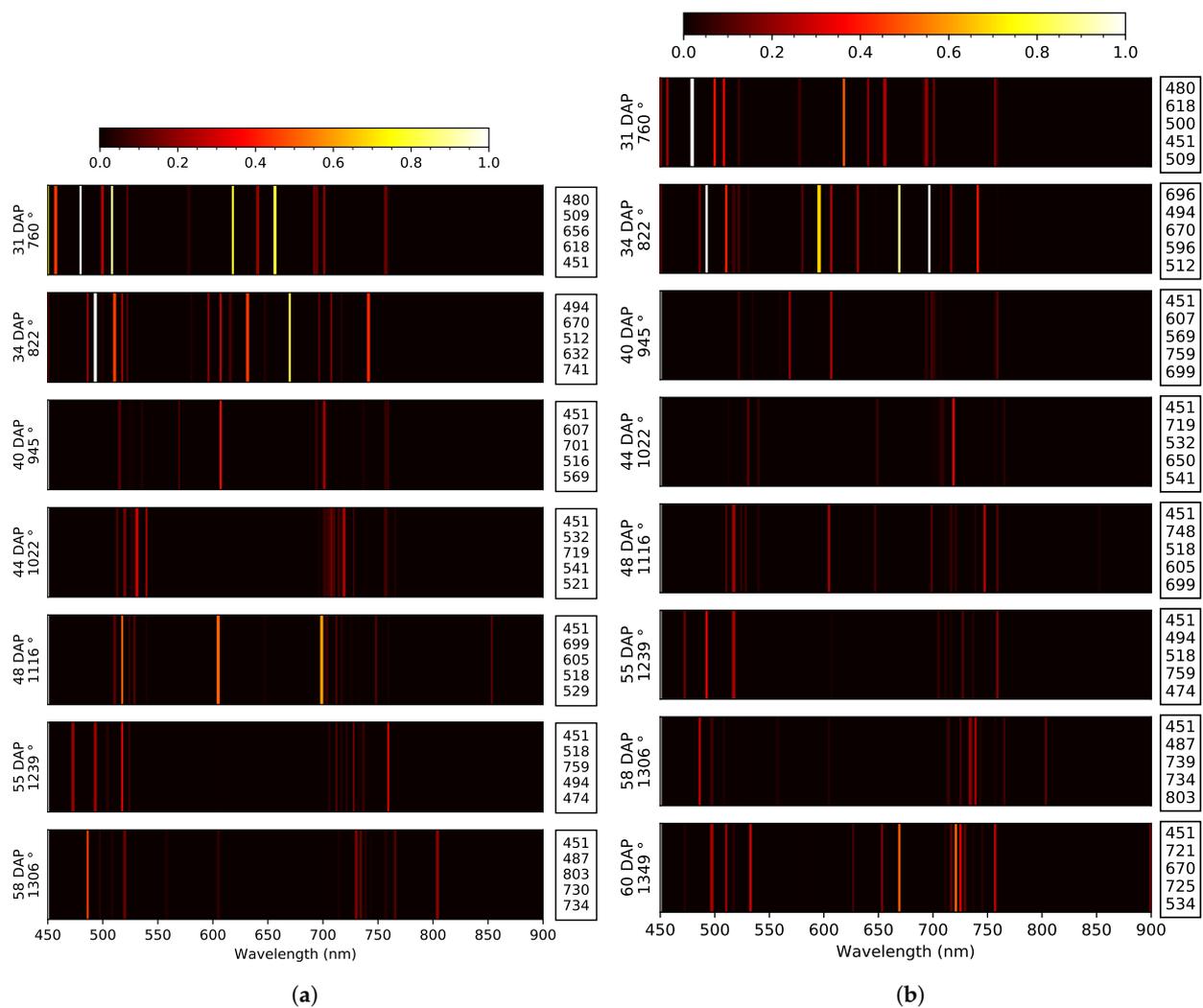


Figure 20. The snap bean 2020 seed length average rankings across for SAM models at (a) the early and (b) the late harvest stages. The corresponding days after planting (DAP) and heat unit for the collected data are shown to the left of each ranking belt, while the top five features, with the first one having the highest score, are listed to the right.

4. Discussion

The variability assessment using PCA resulted in impact plots, as previously depicted in Figure 8. PC-1 reflect changes in 500–600 nm (green) and 600–700 nm (red reflective pigment) with both being coupled to chlorophyll absorption, as well as an observable variability between 700–800 nm (the red edge peak), and the 800–1000 nm spectral region, which represents protein, oil, water, and starch absorption [18]. Results from PC-2, mainly responsible for shifts in amplitude in the 400–700 nm range, address chlorophyll absorption [18], and is consistent with our previously published work in terms of addressing the slope of the vegetation reflectance curve (see [4]).

We noticed that, for the majority of the data sets and models tested, the temporal variability in performance for RENDVI was higher than that of SAM, as well as RENDVI outperforming SAM at earlier growth stages. Further analysis revealed that this difference in performance is due to the fact that RENDVI struggles with differentiating between soil, vegetation shadow, and vegetation in denser crop canopies. Figure 21 shows two plots, each treated with both RENDVI and SAM vegetation detection approaches. The difference between vegetation detection performance for two plots (Figure 21a,b, with Figure 21b being denser than Figure 21a) is evident in that RENDVI performs accurately when there is a distinction between rows of vegetation. As the vegetation grows larger

and denser, RENDVI fails to maintain performance and is not able to discern between vegetation, vegetation shadow, and soil, while SAM consistently does. This translated into shadow/background pixels affecting a plot's average spectrum and thus the final yield estimates. The next consideration relates to the wavelengths required for accurate yield modeling, which has bearing on extending our results to operational sensors with fewer wavelengths (bands).

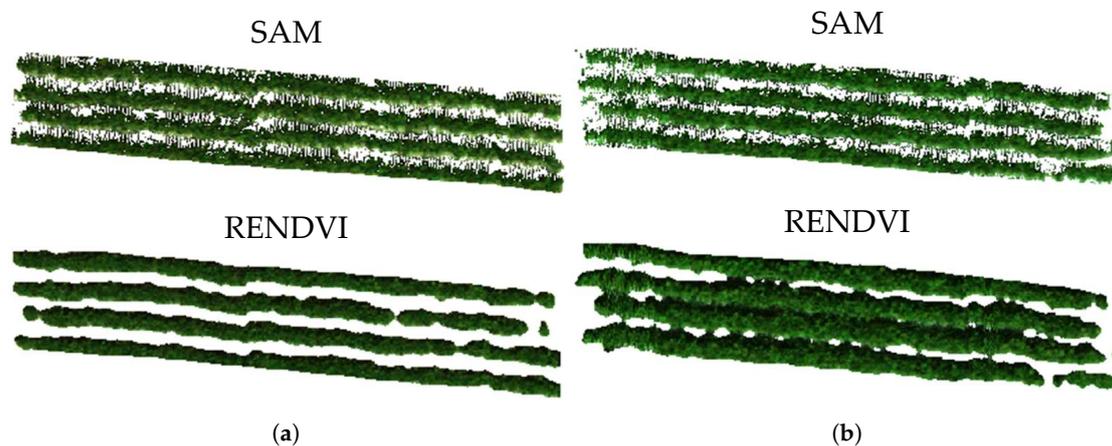


Figure 21. A comparison of the vegetation detection results for cases where (a) vegetation canopy and the background are distinguishable and (b) canopy closure has commenced, i.e., a dense canopy is present.

Table 3 shows a comparison of data sets across harvest stages for pod weight and seed length, specifically (i) the different wavelength regions (zones) and their corresponding dates in DAP (e.g., Blue: E. 50 implies that the blue reflective region is available at all dates, except at 50 DAP), (ii) the equivalent wavelengths detected and the number of times they appeared in the analysis (e.g., 3x: 451 means that 451 nm appeared three times across growth stages), and (iii) when a drop in accuracy (dip) occurred in terms of DAP. The final row also lists similarities between data sets in terms of identified wavelengths. We can see from this table that for pod weight data and 2019 early and late data sets, blue, green, red, and red edge (RE) zones are similar in terms of when they appear during growth. Moreover, the selected wavelengths are highly alike with mutual bands at 451, 505, and 525 nm. The green zone is the most similar and wavelengths at 451, ~520, ~650, and ~760 nm are mutual for both early and late harvest stages in the 2020 pod weight data set. It can be seen that 707 nm was repeated twice for the 2020 pod weight, late harvest stage, something that did not recur for the 2019 data set. In terms of seed length and starting with the 2019 data set, we observed more similarity across spectral regions, but mostly for the red, red edge, and NIR regions. The recurrent wavelengths are highly alike, with mutual wavelengths at 451, 585, ~700, and ~716 nm. Last but not least, the 2020 late harvest stage data set exhibited the highest similarity for all regions, as well as for identified wavelengths. It can be seen that the highest number of recurrences were found for 451 nm, and the rest of the wavelengths reside in 520, 670, 700, and 760 nm. It is important to mention that since the identified wavelengths are decorrelated, it is justified to group wavelengths in close proximity to each other. For example, 707 nm in the 2020 late harvest for pod weight likely would be highly correlated with 699 nm for the 2020 late seed length data. Keeping this in mind, we could confidently say that prominent mutual wavelengths across all data sets (across year and yield indicator type) are ~450, ~500, ~520, ~650, ~710, ~760 nm, with physiological ties to chlorophyll absorption and density, as well as and plant vigor [18,70]. Moreover, the literature states that 450 nm is sensitive to chlorophyll contents $<0.04 \text{ mmol/m}^2$, while also having ties to carotenoid content [71]. The identified wavelength at 705 nm best describes chlorophyll absorption and is highly correlated with this chemical attribute [10,71]. The wavelength at ~520 nm has been linked to the xanthophyll pigment, which could describe the yellowing in large sieve cultivars,

i.e., Venture (L1) and Huntington (L2), as they senesce [72]. If we compare yield indicators with one another, we can see that NIR zones are only apparent for the 2019 late harvest pod weight data set; however, representation from the NIR region can be found in all seed length data sets. Although they did not recur at times, this spectral region could be a predictor of seed length variability (see Figures 17 and 20). As a final note, the identified wavelengths in this study are aligned with those detected in our initial greenhouse yield assessment study (~500, ~550, ~660, and 720–740 nm). This notion shows that detected wavelengths from the greenhouse study to the current UAV study are scalable.

Table 3. Snap bean outperforming models and their corresponding prominent spectral regions with respect to DA (“Zones”) identified distinguishing wavelengths, and the drop in accuracy (“Dip”) in terms of temporal variability.

Data Set	Yield Indicator						Across Yield Indicator WL (nm) Similarity
	Zone (DAP)	Pod Weight WL ^a (nm)	Dip	Zone (DAP)	Seed Length WL (nm)	Dip	
2019 early	Blue: E. ^b 50 Green: T. ^c Red: E. 39, 46 RE: E. 48, 54 NIR: NA ^d	3x: 451 2x: 478, 505, 516, 525	at 39 DAP None	Blue: E. 39, 48 Green: E. 50, 54 Red: T. RE: E. 35, 48 NIR: O. ^e 50, 54	2x: 451, 585, 659, 710, 716	at 40 DAP at 55 DAP	451
2019 late	Blue: E. 35, 50 Green: T. Red: E. 39 RE: E. 46, 54, 56 NIR: O. 35, 56	4x: 525 3x: 451 2x: 505, 541, 659, 759, 819	None at 56 DAP	Blue: E. 35, 50 Green: E. 54, 56 Red: E. 54 RE: O. 48 NIR: O. 54	3x: 451, 721 2x: 583, 585, 659, 716	at 46 DAP at 56 DAP	451, 659
2020 early	Blue: T. Green: E. 44 Red: E. 58 RE: T. NIR: NA	5x: 451, 759 3x: 523 2x: 518, 607, 647	at 34 DAP at 55 DAP	Blue: T. Green: E. 58 Red: E. 44, 55, 58 RE: E. 31, 48 NIR: O. 58	6x: 451 2x: 494, 518	at 44 DAP None	451, 518
2020 late	Blue: E. 58 Green: E. 44 Red: E. 40 RE: E. 31, 55, 60 NIR: NA	4x: 451, 756 2x: 474, 500, 525, 654, 707	at 34 DAP at 60 DAP	Blue: T. Green: E. 58 Red: E. 58, 60 RE: E. 31, 34 NIR: O. 58	7x: 451 2x: 494, 518, 670, 699, 759	at 44 DAP at 60 DAP	451, ~500, ~520, ~700, ~760
Across years WL (nm) similarity		451, ~500, ~520, ~650, ~760			451, ~520, ~500, ~585, ~660, ~720		~451, ~500, ~520, ~650, ~700, ~760

^a WL: Wavelength; ^b E.: Except; ^c T.: Throughout; ^d NA: Not Available; ^e O.: Only.

As previously shown in Figures 9, 12, 15 and 18, and also depicted in Table 3, two time periods show a drop in accuracy for almost all data sets, in terms of temporal performance. The first occurs at the final stages of flowering and the on-set of pod formation, and the second at the final stages of pod formation, with slightly lower performance, for both yield indicators. It can be seen from Table 3 that for pod weight assessment, the initial dip occurs around 34–39 DAP, and the second occurs from 55–60 DAP. In contrast, seed length evaluation, showed that the period at 40–46 DAP accounted for an initial dip, while 56–60 DAP showed a similar drop in performance at the final growth stages. This arguably is one of the remarkable findings of this study, i.e., the window of opportunity between 44 and 55 DAP ensures the highest performance in terms of yield prediction, with an overall ascending trend in accuracy, for both pod weight and seed length yield indicators. Zhao et al. (2007) also observed an increased performance ($R = 0.56–0.89$, for two years of data) in early flowering stages of yield prediction for lint crop [73]. Moreover, Hassan et al. studied wheat yield prediction and observed a decrease in accuracy metrics after the flowering stage [74]. Nothing was found in the literature regarding the latter drop in accuracy. However, since this slight drop in predictive accuracy is usually near or at the day of harvest, it arguably has very little impact on growers, in terms of management impacts. Growers prefer early-season yield assessment that could assist them with proactive managerial and logistical decisions.

Our findings showed that seed length has a stronger focus on the NIR spectral region in terms of identified bands, when compared to pod weight results. While this might be

true for some days, we believe that the incorporation of NIR bands does not necessarily guarantee a higher performance for yield prediction, but rather implies that seed length is more sensitive, as dependent variable, when compared to pod weight, in the case of hyperspectral remotely sensed data. This is a valuable finding for future research, since seed length ground truth data collection is more straightforward in comparison to pod weight, which requires more complex tools and additional human resources. Seed length measurement only requires sampling of ten random pods from the largest pod size, which can be achieved at a lower expense. Finally, as far as the annual data sets are concerned, an increase in performance was observed for the 2020 data set, compared to the 2019 data set, for both pod weight and seed length yield indicators. This increase might be due to the use of 4-point ELM reflectance calibration for the 2020 data set, compared to 3-point ELM for the 2019 data set. This increase in calibration accuracy in the initial stages thus could positively impact the performance of especially the vegetation detection algorithm, but also the overall, average spectral signature of the plot.

5. Conclusions

To the best of our knowledge, this is the first comprehensive study of snap bean yield prediction, early in time during the growth season, via two different yield indicators, pod weight and seed length, with the use of remotely sensed, unmanned aerial system (UAS) hyperspectral data. The objectives of the presented research were to (i) quantify yield prediction of snap bean prior to harvest, with the use of descriptive models, via hyperspectral data, (ii) identify distinguishing spectral features that best predict pod weight and seed length prior to harvest, and (iii) identify the growth/time period that corresponds to the highest predictive accuracies. We first implemented an established hyperspectral noise removal approach to mitigate the impact of noise on the analysis—this approach increased the signal-to-noise ratio (SNR) by 25%. We introduced and implemented a novel feature selection library for Python “Jostar”. We opted to detect only five decorrelated wavelengths, since this number of bands is best suited for transfer to a multispectral, operational sensing system. The proposed ranking method in this work accounted for the effect of rankings from multiple optimization models to garner a holistic understanding of identified features.

Our results show that the pod weight yield indicator could be predicted with high accuracies ($R^2 = 0.78\text{--}0.93$) and low errors ($RMSE = 940\text{--}1369$ kg/ha). Nevertheless, findings from the seed length yield indicator outperformed those of pod weight, with higher accuracies ($R^2 = 0.83\text{--}0.98$) and lower errors ($RMSE = 4.245\text{--}6.018$ mm). Growers rely on both seed length and pod weight as yield indicators to forecast profit and factory feedstock. The performance for seed length yield assessment, along with the simplicity of this type of ground truth data collection, indicates that this is a promising candidate for future work, thus ensuring larger sample sizes and faster and less labor-intensive field work.

Our wavelength assessment revealed that the most recurrent wavelengths, across growth stages and yield indicators, were located in the spectral regions of 450, 500, 520, 650, 700, and 760 nm. We also observed a slight drop in predictive accuracy at the final stages of flowering or at the on-set of pod formation, as well the final stages of pod formation. Our analysis showed that the ideal time period to evaluate yield, via either pod weight or seed length, is 44–55 DAP with an overall ascending trend during this period. Our findings show that a rapid data collection approach, based on a multispectral system that is tuned with the reported wavelengths and aimed to assess snap bean yield in the introduced time period, could potentially be used as a more affordable alternative to the hyperspectral system used in this study. However, the actual implementation of this approach calls for further analysis and mandates future work.

One limitation to this study was the number of plots (sample size) evaluated for yield, which were 18 and 24 for the 2019 and 2020 data sets, respectively. Future work should include larger sample size, since the larger the sample size, the more generalized/robust the model. Another limitation was that the two research sites are both situated in upstate New York, where the weather from one year to another is not significantly variable. Future work

should focus on incorporating fields from different geographical locations, and especially those with varying weather conditions, as well as integrating extreme weather conditions, such as drought and flood. Additionally, since this study was designed to focus on a single year's worth, while a more generalizable model should incorporate predictive abilities across years and data sets. An additional limitation was the high computational cost of the implemented denoising algorithm and running wrapper optimization models for thousands of iterations, while optimizing via the cross-validation set. This resulted in our analysis being limited to one regression model and two vegetation detection approaches. Future work could focus on further integrating vegetation detection methods and various regression models. A possible alternative to our computational approach could be to manipulate the structure of Jostar in order for each agent and its corresponding fitness function to run on a Graphics Processing Unit (GPU), therefore ensuring a significant improvement in speed.

The outcomes of this study bode well for an extension of initial greenhouse results (see [3]) to an airborne, UAS-based approach, without a significant deterioration in accuracies. In fact, the identified wavelengths arguably can be "programmed" into a multispectral, more affordable, and operational sensing platform, while the specific growth stages can be targeted for the most accurate yield predictions. Such a solution would represent a true research-to-application technology transfer and potentially could provide an impetus for the adoption of advanced precision agriculture approaches.

Author Contributions: Conceptualization, A.H., F.Z., J.v.A., S.P.M. and S.J.P.; methodology, A.H., F.Z., J.v.A.; validation, A.H. and J.v.A.; formal analysis, A.H. and J.v.A.; investigation, A.H. and J.v.A.; resources, A.H., F.Z., J.v.A., S.P.M., S.J.P.; Writing—Original Draft preparation, A.H.; Writing—Review and Editing, A.H., F.Z., J.v.A., S.P.M. and S.J.P.; visualization, A.H.; supervision, J.v.A.; project administration, J.v.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science Foundation Partnerships for Innovation (PFI) program (Award Number #1827551).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Nina Raqueno and Tim Bauch, from the center for imaging science's data acquisition team for their time and effort for collecting high quality data for this research study. We also appreciate the reviewers' feedbacks and editor's constructive comments to improve the quality of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACO	ant colony optimization
DAP	days after planting
ELM	empirical line method
FWHM	full-width-half-maximum
GA	genetic algorithm
GDP	gross domestic product
GPU	graphics processing unit
GSD	ground sampling distance
KNN	K-nearest neighbors
LiDAR	light detection and ranging
LRS	plus-L minus-R
MAP	minimum average partial

NIR	near infrared
NSGAI	non-dominated sorting genetic algorithm-II
PA	precision agriculture
PC	principal component (PC)
PCA	principal component analysis
PLSR	partial least squares regression
PSO	particle Swarm Optimization
R^2	coefficient of determination
RE	red edge
RENDVI	red-edge normalized difference vegetation index
RF	random forests
RFE	recursive feature elimination
RMSE	root mean square error
RS	remote sensing
SA	simulated annealing
SAM	spectral angle mapper
SBS	sequential backward search
SFS	sequential forward search
SVR	support vector regression
UAS	unmanned aerial systems
VI	vegetation indices
VIS	visible

References

- Ag and Food Sectors and the Economy. 2020. Available online: <https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/ag-and-food-sectors-and-the-economy> (accessed on 24 January 2021).
- Schellberg, J.; Hill, M.J.; Gerhards, R.; Rothmund, M.; Braun, M. Precision agriculture on grassland: Applications, perspectives and constraints. *Eur. J. Agron.* **2008**, *29*, 59–71. [[CrossRef](#)]
- Hassanzadeh, A.; van Aardt, J.; Murphy, S.P.; Pethybridge, S.J. Yield modeling of snap bean based on hyperspectral sensing: A greenhouse study. *J. Appl. Remote Sens.* **2020**, *14*, 024519. [[CrossRef](#)]
- Hassanzadeh, A.; Murphy, S.P.; Pethybridge, S.J.; van Aardt, J. Growth Stage Classification and Harvest Scheduling of Snap Bean Using Hyperspectral Sensing: A Greenhouse Study. *Remote Sens.* **2020**, *12*, 3809. [[CrossRef](#)]
- Elarab, M.; Toclavilca, A.M.; Torres-Rua, A.F.; Maslova, I.; McKee, M. Estimating chlorophyll with thermal and broadband multispectral high resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *43*, 32–42. [[CrossRef](#)]
- Yu, X.; Lu, H.; Liu, Q. Deep-learning-based regression model and hyperspectral imaging for rapid detection of nitrogen concentration in oilseed rape (*Brassica napus* L.) leaf. *Chemom. Intell. Lab. Syst.* **2018**, *172*, 188–193. [[CrossRef](#)]
- Tilling, A.K.; O’Leary, G.J.; Ferwerda, J.G.; Jones, S.D.; Fitzgerald, G.J.; Rodriguez, D.; Belford, R. Remote sensing of nitrogen and water stress in wheat. *Field Crop. Res.* **2007**, *104*, 77–85. [[CrossRef](#)]
- Thorp, K.; Tian, L. A review on remote sensing of weeds in agriculture. *Precis. Agric.* **2004**, *5*, 477–508. [[CrossRef](#)]
- Christy, C.D. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Comput. Electron. Agric.* **2008**, *61*, 10–19. [[CrossRef](#)]
- Lleó, L.; Barreiro, P.; Ruiz-Altisent, M.; Herrero, A. Multispectral images of peach related to firmness and maturity at harvest. *J. Food Eng.* **2009**, *93*, 229–235. [[CrossRef](#)]
- Lee, W.S.; Alchanatis, V.; Yang, C.; Hirafuji, M.; Moshou, D.; Li, C. Sensing technologies for precision specialty crop production. *Comput. Electron. Agric.* **2010**, *74*, 2–33. [[CrossRef](#)]
- Mulla, D.J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst. Eng.* **2013**, *114*, 358–371. [[CrossRef](#)]
- Schwalbert, R.A.; Amado, T.; Corassa, G.; Pott, L.P.; Prasad, P.V.; Ciampitti, I.A. Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. *Agric. For. Meteorol.* **2020**, *284*, 107886. [[CrossRef](#)]
- Li, B.; Xu, X.; Zhang, L.; Han, J.; Bian, C.; Li, G.; Liu, J.; Jin, L. Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 161–172. [[CrossRef](#)]
- Feng, L.; Zhang, Z.; Ma, Y.; Du, Q.; Williams, P.; Drewry, J.; Luck, B. Alfalfa Yield Prediction Using UAV-Based Hyperspectral Imagery and Ensemble Learning. *Remote Sens.* **2020**, *12*, 2028. [[CrossRef](#)]
- Nasrabi, N.M. Hyperspectral target detection: An overview of current and future challenges. *IEEE Signal Process. Mag.* **2013**, *31*, 34–44. [[CrossRef](#)]
- Im, J.; Jensen, J.R. Hyperspectral remote sensing of vegetation. *Geogr. Compass* **2008**, *2*, 1943–1961. [[CrossRef](#)]
- Curran, P.J. Remote sensing of foliar chemistry. *Remote Sens. Environ.* **1989**, *30*, 271–278. [[CrossRef](#)]

19. Plaza, J.; Plaza, A.; Perez, R.; Martinez, P. On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images. *Pattern Recognition* **2009**, *42*, 3032–3045. [CrossRef]
20. Rasti, B.; Scheunders, P.; Ghamisi, P.; Licciardi, G.; Chanussot, J. Noise reduction in hyperspectral imagery: Overview and application. *Remote Sens.* **2018**, *10*, 482. [CrossRef]
21. Yuan, Q.; Zhang, L.; Shen, H. Hyperspectral image denoising employing a spectral–spatial adaptive total variation model. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3660–3677. [CrossRef]
22. Kerekes, J.P.; Baum, J.E. Hyperspectral imaging system modeling. *Linc. Lab. J.* **2003**, *14*, 117–130. Available online: https://www.cis.rit.edu/people/faculty/kerekes/pdfs/LLJ_2003_Kerekes.pdf (accessed on 24 January 2021).
23. Griffin, M.K.; Burke, H.H.K. Compensation of hyperspectral data for atmospheric effects. *Linc. Lab. J.* **2003**, *14*, 29–54. Available online: https://129.55.110.66/publications/journal/pdf/vol14_no1/14_1compensation.pdf (accessed on 24 January 2021).
24. Antczak, K. Deep recurrent neural networks for ECG signal denoising. *arXiv* **2018**, arXiv:1807.11551. Available online: <https://arxiv.org/abs/1807.11551v3> (accessed on 24 January 2021).
25. Jiang, Y.; Li, H.; Rangaswamy, M. Deep learning denoising based line spectral estimation. *IEEE Signal Process. Lett.* **2019**, *26*, 1573–1577. [CrossRef]
26. Zhang, C.; Zhou, L.; Zhao, Y.; Zhu, S.; Liu, F.; He, Y. Noise reduction in the spectral domain of hyperspectral images using denoising autoencoder methods. *Chemom. Intell. Lab. Syst.* **2020**, *203*, 104063. [CrossRef]
27. Schulze, G.; Jirasek, A.; Marcia, M.; Lim, A.; Turner, R.F.; Blades, M.W. Investigation of selected baseline removal techniques as candidates for automated implementation. *Appl. Spectrosc.* **2005**, *59*, 545–574. [CrossRef]
28. Chen, Y.; Oliver, D.S.; Zhang, D. Data assimilation for nonlinear problems by ensemble Kalman filter with reparameterization. *J. Pet. Sci. Eng.* **2009**, *66*, 1–14. [CrossRef]
29. Fernández-Navales, J.; López, M.I.; Sánchez, M.T.; Morales, J.; González-Caballero, V. Shortwave-near infrared spectroscopy for determination of reducing sugar content during grape ripening, winemaking, and aging of white and red wines. *Food Res. Int.* **2009**, *42*, 285–291. [CrossRef]
30. Lang, M.; Guo, H.; Odegard, J.E.; Burrus, C.S.; Wells, R.O. Noise reduction using an undecimated discrete wavelet transform. *IEEE Signal Process. Lett.* **1996**, *3*, 10–12. [CrossRef]
31. Jansen, M. *Noise Reduction by Wavelet Thresholding*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012, Volume 161. [CrossRef]
32. Chen, G.; Qian, S.E. Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. *IEEE Trans. Geosci. Remote Sens.* **2010**, *49*, 973–980. [CrossRef]
33. Zwick, W.R.; Velicer, W.F. Comparison of five rules for determining the number of components to retain. *Psychol. Bull.* **1986**, *99*, 432. Available online: <https://files.eric.ed.gov/fulltext/ED251510.pdf> (accessed on 24 January 2021). [CrossRef]
34. Bartlett, M.S. Tests of significance in factor analysis. *Br. J. Stat. Psychol.* **1950**, *3*, 77–85. [CrossRef]
35. Kaiser, H.F. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* **1960**, *20*, 141–151. [CrossRef]
36. Velicer, W.F. The relation between factor score estimates, image scores, and principal component scores. *Educ. Psychol. Meas.* **1976**, *36*, 149–159. [CrossRef]
37. Cattell, R.B. The scree test for the number of factors. *Multivar. Behav. Res.* **1966**, *1*, 245–276. [CrossRef] [PubMed]
38. Horn, J.L. A rationale and test for the number of factors in factor analysis. *Psychometrika* **1965**, *30*, 179–185. [CrossRef] [PubMed]
39. Hang, R.; Liu, Q.; Hong, D.; Ghamisi, P. Cascaded recurrent neural networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5384–5394. [CrossRef]
40. Agarwal, A.; El-Ghazawi, T.; El-Askary, H.; Le-Moigne, J. Efficient hierarchical-PCA dimension reduction for hyperspectral imagery. In Proceedings of the 2007 IEEE International Symposium on Signal Processing and Information Technology, Giza, Egypt, 15–18 December 2007; pp. 353–356. [CrossRef]
41. Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–4. [CrossRef]
42. Hawkins, D.M. The problem of overfitting. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1–12. [CrossRef] [PubMed]
43. Venkatesh, B.; Anuradha, J. A review of feature selection and its methods. *Cybern. Inf. Technol.* **2019**, *19*, 3–26. [CrossRef]
44. Pearson, K. LIII. On lines and planes of closest fit to systems of points in space. *London Edinburgh Dublin Philos. Mag. J. Sci.* **1901**, *2*, 559–572. [CrossRef]
45. Battiti, R. Using mutual information for selecting features in supervised neural net learning. *IEEE Trans. Neural Netw.* **1994**, *5*, 537–550. [CrossRef]
46. Kumar, V.; Minz, S. Feature selection: A literature review. *SmartCR* **2014**, *4*, 211–229. [CrossRef]
47. Goldenberg, D.E. Genetic Algorithms in Search, Optimization and Machine Learning. 1989. Available online: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/46947/10994_2005_Article_422926.pdf (accessed on 24 January 2021)
48. Li, B.; Wang, L.; Song, W. Ant colony optimization for the traveling salesman problem based on ants with memory. In Proceedings of the 2008 Fourth International Conference on Natural Computation, Jinan, China, 18–20 October 2008, Volume 7, pp. 496–501. [CrossRef]
49. Del Valle, Y.; Venayagamoorthy, G.K.; Mohagheghi, S.; Hernandez, J.C.; Harley, R.G. Particle swarm optimization: Basic concepts, variants and applications in power systems. *IEEE Trans. Evol. Comput.* **2008**, *12*, 171–195. [CrossRef]

50. John, G.H.; Kohavi, R.; Pflieger, K. Irrelevant features and the subset selection problem. In *Machine Learning Proceedings 1994*; Elsevier: Amsterdam, The Netherlands, 1994; pp. 121–129. [CrossRef]
51. Ma, S.; Huang, J. Penalized feature selection and classification in bioinformatics. *Briefings Bioinform.* **2008**, *9*, 392–403. [CrossRef]
52. Tao, H.; Feng, H.; Xu, L.; Miao, M.; Yang, G.; Yang, X.; Fan, L. Estimation of the yield and plant height of winter wheat using UAV-based hyperspectral images. *Sensors* **2020**, *20*, 1231. [CrossRef]
53. USDA. *USDA National Agricultural Statistics Service*; National Statistical Services. 2016. Available online: https://quickstats.nass.usda.gov/results/5B4DC16F-4537-38BB-86AC-010B3FBF83DC?pivot=short_desc (accessed on 24 January 2021).
54. Yuan, M.; Ruark, M.D.; Bland, W.L. A simple model for snap bean (*Phaseolus vulgaris* L.) development, growth and yield in response to nitrogen. *Field Crop. Res.* **2017**, *211*, 125–136. [CrossRef]
55. Saleh, S.; Liu, G.; Liu, M.; Ji, Y.; He, H.; Gruda, N. Effect of Irrigation on Growth, Yield, and Chemical Composition of Two Green Bean Cultivars. *Horticulturae* **2018**, *4*, 3. [CrossRef]
56. Kaputa, D.S.; Bauch, T.; Roberts, C.; McKeown, D.; Foote, M.; Salvaggio, C. Mx-1: A new multi-modal remote sensing uas payload with high accuracy gps and imu. In Proceedings of the 2019 IEEE Systems and Technologies for Remote Sensing Applications Through Unmanned Aerial Systems (STRATUS), Rochester, NY, USA, 25–27 February 2019; pp. 1–4. [CrossRef]
57. Mamaghani, B.; Salvaggio, C. Comparative study of panel and panelless-based reflectance conversion techniques for agricultural remote sensing. *arXiv* **2019**, arXiv:1910.03734. Available online: <https://arxiv.org/abs/1910.03734> (accessed on 24 January 2021).
58. Lee, S.U.; Chung, S.Y.; Park, R.H. A comparative performance study of several global thresholding techniques for segmentation. *Comput. Vision Graph. Image Process.* **1990**, *52*, 171–190. [CrossRef]
59. Yuhas, R.H.; Goetz, A.F.; Boardman, J.W. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In Proceedings of the Summaries 3rd Annual JPL Airborne Geoscience Workshop, Pasadena, CA, USA, 1–5 June 1992; Volume 1, pp. 147–149
60. Garrido, L.E.; Abad, F.J.; Ponsoda, V. Performance of Velicer’s minimum average partial factor retention method with categorical variables. *Educ. Psychol. Meas.* **2011**, *71*, 551–570. [CrossRef]
61. Price, K.V. Differential evolution. In *Handbook of Optimization*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 187–214. [CrossRef]
62. Van Laarhoven, P.J.; Aarts, E.H. Simulated annealing. In *Simulated Annealing: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 1987; pp. 7–15. [CrossRef]
63. Levin, L.A. Universal sequential search problems. *Probl. Peredachi Inform.* **1973**, *9*, 115–116. Available online: <http://www.mathnet.ru/links/e8c7e04918539cbbc30a0f5ff178abdc/ppi914.pdf> (accessed on 24 January 2021).
64. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]
65. Altmann, A.; Toloşi, L.; Sander, O.; Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinformatics* **2010**, *26*, 1340–1347. [CrossRef]
66. Van Rossum, G.; Drake, F.L., Jr. *Python Tutorial*; Centrum voor Wiskunde en Informatica: Amsterdam, The Netherlands, 1995. Available online: http://lib.21h.io/library/G7B3RFY7/download/GBMFTU3C/Python_3.8.3_Docs_.pdf (accessed on 24 January 2021).
67. Bergstra, J.; Yamins, D.; Cox, D.D. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In Proceedings of the 12th Python in Science Conference, Citeseer, Austin, TX, USA, 12–18 July 2013; Volume 13, p. 20. Available online: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.704.3494&rep=rep1&type=pdf> (accessed on 24 January 2021).
68. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
69. Bajorski, P. *Statistics for Imaging, Optics, and Photonics*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 808. [CrossRef]
70. Wei, F.; Yan, Z.; Yongchao, T.; Weixing, C.; Xia, Y.; Yingxue, L. Monitoring leaf nitrogen accumulation in wheat with hyper-spectral remote sensing. *Acta Ecol. Sin.* **2008**, *28*, 23–32. [CrossRef]
71. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [CrossRef]
72. Adão, T.; Hruška, J.; Pádua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J.J. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sens.* **2017**, *9*, 1110. [CrossRef]
73. Zhao, D.; Reddy, K.R.; Kakani, V.G.; Read, J.J.; Koti, S. Canopy reflectance in cotton for growth assessment and lint yield prediction. *Eur. J. Agron.* **2007**, *26*, 335–344. [CrossRef]
74. Hassan, M.A.; Yang, M.; Rasheed, A.; Yang, G.; Reynolds, M.; Xia, X.; Xiao, Y.; He, Z. A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Sci.* **2019**, *282*, 95–103. [CrossRef]