



Article

Alfalfa Yield Prediction Using UAV-Based Hyperspectral Imagery and Ensemble Learning

Luwei Feng ^{1,2}, Zhou Zhang ^{1,*} , Yuchi Ma ¹, Qingyun Du ^{2,3,4} , Parker Williams ¹, Jessica Drewry ¹  and Brian Luck ¹

¹ Biological Systems Engineering, University of Wisconsin–Madison, Madison, WI 53706, USA; lwfeng@whu.edu.cn (L.F.); ma286@wisc.edu (Y.M.); pjwilliams3@wisc.edu (P.W.); jdrewry@wisc.edu (J.D.); bluck@wisc.edu (B.L.)

² School of Resources and Environmental Science, Wuhan University, Wuhan 430079, China; qydu@whu.edu.cn

³ Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University, Wuhan 430079, China

⁴ Key Laboratory of Digital Mapping and Land Information Application Engineering, Ministry of Natural Resources, Wuhan University, Wuhan 430079, China

* Correspondence: zzhang347@wisc.edu

Received: 3 June 2020; Accepted: 23 June 2020; Published: 24 June 2020



Abstract: Alfalfa is a valuable and intensively produced forage crop in the United States, and the timely estimation of its yield can inform precision management decisions. However, traditional yield assessment approaches are laborious and time-consuming, and thus hinder the acquisition of timely information at the field scale. Recently, unmanned aerial vehicles (UAVs) have gained significant attention in precision agriculture due to their efficiency in data acquisition. In addition, compared with other imaging modalities, hyperspectral data can offer higher spectral fidelity for constructing narrow-band vegetation indices which are of great importance in yield modeling. In this study, we performed an in-season alfalfa yield prediction using UAV-based hyperspectral images. Specifically, we firstly extracted a large number of hyperspectral indices from the original data and performed a feature selection to reduce the data dimensionality. Then, an ensemble machine learning model was developed by combining three widely used base learners including random forest (RF), support vector regression (SVR) and K-nearest neighbors (KNN). The model performance was evaluated on experimental fields in Wisconsin. Our results showed that the ensemble model outperformed all the base learners and a coefficient of determination (R^2) of 0.874 was achieved when using the selected features. In addition, we also evaluated the model adaptability on different machinery compaction treatments, and the results further demonstrate the efficacy of the proposed ensemble model.

Keywords: alfalfa; yield prediction; hyperspectral; unmanned aerial vehicle (UAV); ensemble learning; vegetation index

1. Introduction

Alfalfa is one of the most important and widespread perennial legumes, and it is considered as a valuable forage crop with relatively high yield and nutritional value [1]. In 2018, nearly 53 million tons of alfalfa and alfalfa mixtures were harvested from about 17 million acres in the United States [2]. Due to its large scale, precise management to achieve the forage yield goal is critical to optimize profitability [3]. Timely estimation of alfalfa production within the growing season can inform precision management decisions to reduce the potential production loss. Additionally, rapidly

and accurately estimating the yield within the growing season also has the potential to improve the timing in harvesting alfalfa to optimize the forage quality and production [4].

The early development of alfalfa is characterized by the formulation and growth of leaves and stems. Then the appearance of flower buds indicates the transformation from vegetative stage to reproductive stage, after which alfalfa progressively passes through flowering and seed pod development [5]. At different stages of an alfalfa life cycle, adequate fertilization, irrigation and tillage practices are required to ensure long-lasting productive stands [1]. In addition, the cutting frequency also has a strong impact on the alfalfa yield, and three or four cuttings are usually adopted in the northern United States. Typically, decreased yield can be observed in the successive cuttings through a year [6]. Generally, alfalfa is harvested as hay or silage. Alfalfa intended for hay is sun-dried in the field after cutting to a moisture level of less than 12% [7], while shorter drying time is needed for silage which has more than 50% moisture content [8]. Machinery is typically used for alfalfa harvest from mowing, raking, merging, to baling or chopping [9]. Wheel traffic from these machines has a significant impact on soil health and crop production potential. According to a multi-state study, the yield reduction in next cutting caused by wheel traffic ranged from 5% to 26% depending on the traffic timing [10]. Traffic events which occurred three to five days after mowing caused significant losses in alfalfa yield [11]. Though yield loss from wheel traffic varies from field to field, it can be reduced by minimizing the number and locations of machinery operations [12–15].

Traditional yield assessment is based on destructive sampling which is to manually collect data samples in the field and weigh the samples to determine the yield [16–19]. However, destructive sampling is not only laborious and time-consuming, but also are unable to monitor the crop status over a growing season [4]. The advances of remote sensing techniques have provided a non-destructive and efficient way to monitor the crop growth, and thus has great potential for crop yield analysis. In the last few decades, satellite remote sensing has been widely used in agriculture [20–22]. Among these studies, several researchers explored the utility of satellite data to estimate alfalfa yield. For example, multiple vegetation indices (VIs) extracted from time-series Landsat images were used to predict alfalfa yield in Saudi Arabia, and near-infrared (NIR) reflectance, soil adjusted vegetation index (SAVI) and normalized difference vegetation index (NDVI) were found to be strongly correlated with the yield [23]. High-resolution commercial QuickBird satellite data were applied to estimate alfalfa yield over hilly areas on Loess Plateau of China, and better performance was achieved than using the Landsat data [20]. Though successful, the adoption of satellite remote sensing in precision farming has been restricted due to limitations, such as cloud contamination and relatively coarse spatial and temporal resolutions [21,22].

Recently, unmanned aerial vehicles (UAVs) have gained significant attention due to their greater flexibility in mission scheduling, and image data with finer spatial resolution acquired from different sensors mounted on UAV platforms have been widely used in precision agriculture. Based on low-cost conventional digital RGB images, various studies have been carried out to assess the growth status of crops and predict yields using either the original three color bands or derived color indices, such as green-red vegetation index (GRVI) [24] and excess green vegetation index [25]. Since plants typically have strong reflective properties in the NIR wavelengths, multispectral sensors have become more favored by incorporating the NIR channels. For example, several VIs which were developed by including the NIR band, such as NDVI and red edge position index (REPI), have been successfully applied in grain yield prediction [26], crop senescence rate detection [27], plant water status assessment [28], and other applications in precision agriculture [29–31].

Hyperspectral cameras are more expensive than RGB and multispectral cameras, and large data storage capacity is required to store hyperspectral data cubes and perform data processing. Despite the challenges, hyperspectral imagery is capable of providing more detailed spectral information and offering better opportunities for applications in precision agriculture since hundreds of spectral bands, arranged in narrower bandwidth, are consisted in the images. Based on the advantage of hyperspectral imagery, various agricultural applications have been studied using hyperspectral imagery, such as

crop mapping [32,33], disease detection [34,35], and stress assessment [36,37]. Among the wide range of applications, crop yield prediction by taking advantage of the continuous narrow spectral bands, is one of the most popular [38–43]. For example, Leaf area index (LAI) and chlorophyll (CHL) were estimated using noisy-reduced UAV-based hyperspectral data, and these traits were then used to predict wheat yield in Belm, northwestern Germany, achieving a coefficient of determination (R^2) of 0.88 [44]. Another example showed that wheat yield could also be effectively predicted using 190 narrow bands from UAV-based hyperspectral imagery before harvest in Minnesota [45]. Compared to RGB data, hyperspectral datasets were demonstrated to be more effective in modeling the yield of winter barley [43]. Besides, they were also compared to multispectral images in grain sorghum yield prediction, and the narrow-band NDVIs extracted from the hyperspectral data were able to explain more yield variability [42]. Recently, to facilitate the feature selection, a feature explorer interactive system was developed to accelerate the exploration, ranking and selection of the derived hyperspectral indices [46]. Though successful, these studies were all focused on grain crops. It is necessary to investigate the potential of applying hyperspectral imaging for modeling the forage yield.

In general, two types of models have been investigated for modeling the yield: process-based crop simulation models and machine learning models. The crop models forecast yield by simulating the crop growth, based on the known physiological characteristics of plants and a number of environmental factors, and representative alfalfa simulation models include SIMED [47], ALSIM [48], ALFALFA 1.4 [49], ALF2LP [50], and the DSSAT-CROPGRO-Perennial Forage model, which can be adapted to alfalfa [51]. Although successful, these models typically require a large number of input data related to crop cultivar, management practices and soil conditions which are often difficult to obtain [52]. Moreover, the calibration of these mechanistic models can be challenging due to the complexity of physiological processes.

In contrast, machine learning approaches aim to develop an empirical relationship between the independent variables with the yield, and thus have the advantage of forecasting the yield without relying on the specific parameters for individual crops [53]. In this context, various machine learning models have been developed for crop yield prediction, such as linear regression [54], support vector regression (SVR) [55] and artificial neural networks [56]. However, these approaches all depend on single predictive models, and are subject to overfitting with limited training data [57]. In the machine learning community, there is an increasing interest in ensemble techniques, which use a group of base learners for training and combine the predictions from all of them for final predictions. Ensemble models usually result in better predictive performance compared to single models [58,59]. Bagging, boosting and stacking are three commonly used ensemble learning strategies. Bagging generates base learners in parallel using training subsets obtained from bootstrap sampling, and boosting trains a sequence of base models by exploiting their dependences [60]. Random forest (RF) and stochastic gradient boosting are representative bagging and boosting approaches respectively, and their superior performance in yield prediction has been demonstrated in several studies [53,61]. Unlike bagging and boosting which typically combine homogeneous learners, stacking tends to employ heterogeneous learners and leverage differences between them to improve the final accuracy. The diversity condition guarantees the complementary information provided by difference models, which is also the key for ensuring that the incorrect results are unlikely to achieve by all the base learners. To the best of our knowledge, the stacking-based ensemble strategy has not been applied in yield prediction, though several successful applications have been observed in machine learning and computer vision [62,63].

This study was designed to conduct in-season alfalfa yield prediction using UAV-based hyperspectral images. The specific objectives included (1) investigate the potential of using hyperspectral images for alfalfa yield prediction, (2) establish an ensemble learning model to improve prediction performance, and (3) evaluate the model adaptability under different machinery compaction treatments.

2. Materials and Methods

2.1. Experimental Design and Field Data Collection

The research trials were conducted at the Arlington Agricultural Research Station in Wisconsin in 2019 (Figure 1). Wisconsin is located in the Great Lakes region which is in the north-central part of the United States, and Arlington is located in south-central Wisconsin. Arlington has a continental climate characterized by warm and humid summers and cold and dry winters. In 2019, the annual total precipitation reached 115 cm, and the monthly average temperature was highest at 23 °C in July and lowest at −9 °C in April [64].

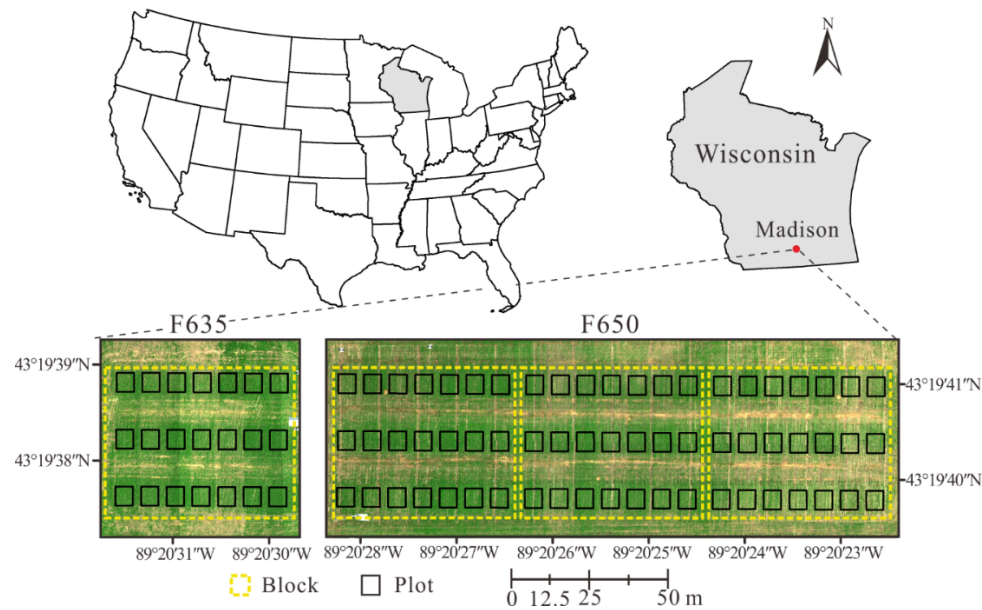


Figure 1. Experimental site location.

The study area shown in Figure 1 consists of two adjacent alfalfa fields, namely F635 and F650. There was one block in F635 where alfalfa was seeded in May, 2018, and three blocks in F650 with alfalfa newly planted in May, 2019. Each block contained 21 plots with a size of 24 m² (W-E 4 m × N-S 6 m), and it was randomly designed with seven levels of machinery traffic treatments and three replications for each. In each treatment, a certain number of traffic passes were used, with each pass simulating a particular machinery operation, such as mowing, raking, merging, baling, chopping and transporting. For treatments T1–T6, a swather was used to simulate both silage and hay harvest by applying weight and pressure to the plant and soil. For example, three passes applied in T2 and T5 were used to simulate mowing, merging (or raking) and harvest operation, and the delayed second and third passes in T5 were because more time was required for alfalfa to dry in hay harvest. The T7 was a control group without any traffic applied. The detailed descriptions for the seven treatments are presented in Table 1. The two alfalfa fields were harvested four times from June to September, and the dry matter yield data acquired and processed from the last two harvests which occurred in August and September were used in this study.

Table 1. Machinery traffic treatments applied to alfalfa plots in 2019.

Treatment	Name	Simulated Traffic	Description
T1	Single Pass Silage/Hay	Mower	One application of compaction immediately after harvest covering the entire plot.
T2	Three Passes Silage	Mower, merger, forage harvester	Three applications of compaction. One immediately after harvest, one 24 h after harvest, and one 26 h after harvest. Full plot application.
T3	Five Passes Silage	Mower, merger, rake, forage harvester, transport vehicle	Five applications of compaction. One immediately after harvest, two passes 24 h after harvest, and two passes 26 h after harvest. Full plot application.
T4	Simulated Silage producer	Mower, merger or rake, forage harvester, transport vehicle	Two-wheel tracks applied within the plot. One pass immediately after harvest, one pass 24 h after harvest, and two passes 26 h after harvest.
T5	Three Passes Hay	Mower, merger or rake, bailer	Three applications of compaction. One immediately after harvest, one 48 h after harvest, and one 72 h after harvest. Full plot application.
T6	Five Passes Hay	Mower, merger, rake, bailer, transport vehicle	Five applications of compaction. One immediately after harvest, two passes 48 h after harvest, and two passes 72 h after harvest. Full plot application.
T7	Zero Passes	No	No machine traffic applied.

2.2. Hyperspectral Image Acquisition and Pre-Processing

The hyperspectral data were acquired by a Headwall nano-hyperspec push-broom scanner. This sensor covers 273 spectral bands ranging between 400–1000 nm with a bandwidth of 2.2 nm [65]. Each scan line contains 640 pixels with a pixel pitch of 7.4 μm . A VectorNav (VN)–300 GNSS/INS navigation system was integrated with the hyperspectral camera to directly provide the position and orientation of the camera for data geo-referencing. The DJI Matrice 600 Pro (M600) was used as the UAV platform. Furthermore, a DJI Ronin MX three-axis gimbal stabilizer was used on the M600 airframe. By using the gimbal, the sensor can maintain a nadir view regardless of airframe orientation. This capability can help stabilize the hyperspectral camera during the flight, leading to improved data geometry quality. Two UAV flights were conducted on July 25 and August 19 in 2019 under cloudless weather conditions. The UAV was flying at a speed of 5 m/s from an altitude of 40 m, and the corresponding ground sampling distance (GSD) was 2.5 cm.

The acquired hyperspectral data were geometrically and radiometrically corrected. For the geometric correction, the data were orthorectified using the Headwall SpectralView software based on the GNSS/INS data from the VN-300. The raw hyperspectral data were first converted to radiance using SpectralView then calibrated to reflectance using the calibration panels with 56%, 32% and 11% reflectivity. Within each plot, the background (e.g., shadow and soil) were removed by thresholding the NIR band at 800 nm wavelength (the pixels with reflectance values below 30% were removed in this study). The filtering strategy was adopted because vegetation typically has much higher reflectance values in the NIR region than the background [66,67]. Similar to previous studies [45,68], the 30% threshold was empirically determined. The alfalfa reflectance data were extracted for each plot and noisy bands below 442 nm and beyond 957 nm were removed.

2.3. Spectral Feature Extraction and Reduction

The acquired hyperspectral data contain hundreds of spectral bands, and many adjacent bands are highly correlated. To reduce the data dependency, instead of using all the original bands, we extracted the narrow-band indices as spectral features and used them for modeling the alfalfa yield in this study. Specifically, we explored 80 published indices (Table 2), with each derived from two or more spectral bands. The indices included simple ratio index (SRI), NDVI, REPI, chlorophyll absorption ratio index, modified versions of these indices such as mND₇₀₅ and combination of them such as TCARI/OSAVI₁. Although the calculations are varied, most wavebands used were in the red and NIR ranges.

Table 2. Summary of vegetation indices (Vis) explored in this study.

Full Form	Index	Formula	Reference
Normalized difference vegetation index	NDVI _[471,584]	$(R584 - R471)/(R584 + R471)$	[69]
	NDVI _[521,689]	$(R689 - R521)/(R689 + R521)$	[69]
	NDVI _[550,760]	$(R760 - R550)/(R760 + R550)$	[70]
	NDVI _[667,740]	$(R740 - R667)/(R740 + R667)$	[71]
	NDVI _[670,800]	$(R800 - R670)/(R800 + R670)$	[72]
	NDVI _[705,750]	$(R750 - R705)/(R750 + R705)$	[73]
	NDVI _[710,750]	$(R750 - R710)/(R750 + R710)$	[74]
	NDVI _[710,780]	$(R780 - R710)/(R780 + R710)$	[75]
	NDVI _[717,732]	$(R750 - R710)/(R750 + R710)$	[76]
	NDVI _[717,770]	$(R732 - R717)/(R732 + R717)$	[76]
Physiological reflectance index	NDVI _[720,820]	$(R820 - R720)/(R820 + R720)$	[77]
	NDVI _[734,750]	$(R750 - R735)/(R750 + R734)$	[76]
Physiological reflectance index	PRI _[528,567]	$(R528 - R567)/(R528 + R567)$	[78]
	PRI _[531,570]	$(R570 - R531)/(R531 + R570)$	[79]
Normalized difference red edge	NDRE	$(R790 - R720)/(R790 + R720)$	[80]
Modified normalized difference vegetation index	mND ₇₀₅	$(R750 - R705)/(R750 + R705 - 2R445)$	[81]
Green normalized difference vegetation index	GNDVI	$(R750 - R550)/(R750 + R550)$	[82]
Renormalized difference vegetation index	RDVI	$(R800 - R670) / \sqrt{R800 + R670}$	[83]
Normalized difference cloud index	NDCI	$(R762 - R527)/(R762 + R527)$	[84]
Curvature index	CI	$R675 \times R690/R6832$	[85]
-	Datt ₁	$(R850 - R710)/(R850 - R680)$	[86]
	Datt ₂	$R850/R710$	
	Datt ₃	$R754/R704$	
Double Difference index	DD	$(R749 - R720) - (R701 - R672)$	[87]
Double peak canopy nitrogen index	DCNI	$(R720 - R700)/[(R700 - R670)(R720 - R670 + 0.03)]$	[88]
-	Gitelson ₁	$1/R700$	[89]
	Gitelson ₂	$(R750-R800/R695-R740) - 1$	[90]
-	Carte ₁	$R695/R760$	[91]
	Carte ₂	$R605/R760$	
	Carte ₃	$R710/R760$	
	Carte ₄	$R695/R670$	
Simple ratio index	SRI _[533,565]	$R565/R533$	[92]
	SRI _[550,750]	$R750/R550$	[93]
	SRI _[550,760]	$R760/R550$	[70]
	SRI _[560,810]	$R810/R560$	[94]
	SRI _[629,734]	$R734/R629$	[71]
	SRI _[660,810]	$R810/R660$	[95]
	SRI _[670,700]	$R700/R670$	[96]

Table 2. Cont.

Full Form	Index	Formula	Reference
Simple ratio index	SRI _[533,565]	R565/R533	[92]
	SRI _[670,800]	R800/R670	[88]
	SRI _[675,700]	R675/R700	[97]
	SRI _[680,800]	R800/R680	[81]
	SRI _[690,752]	R752/R690	[93]
	SRI _[700,750]	R750/R700	[93]
	SRI _[705,750]	R750/R705	[73]
	SRI _[706,755]	R706/R755	[76]
	SRI _[708,747]	R747/R708	[98]
	SRI _[710,750]	R750/R710	[99]
	SRI _[717,741]	R741/R717	[98]
Modified simple ratio index	SRI _[720,735]	R735/R720	[98]
	SRI _[720,738]	R738/R720	[98]
	mSRI _[550,780]	R780/R550-1	[100]
	mSRI _[710,780]	R780/R710-1	[101]
	mSRI _[720,750]	R750/R720-1	[100]
New vegetation index	mSR705	$(R750 - R445)/(R705 - R445)$	[86]
	mSR	$(R750/R705 - 1) / (\sqrt{R750/R705} + 1)$	[102]
	NVI ₁	$(R777 - R747)/R673$	[103]
	NVI ₂	$R705/(R717 + R491)$	[92]
Enhanced vegetation index	EVI	$2.5(R800 - R670)/(R800 - 6R670 - 7.5R475 + 1)$	[104]
Transformed Chlorophyll absorption in reflectance index	TCARI ₁	$3[(R700 - R670) - 0.2(R700 - R550)(R700/R670)]$	[105]
	TCARI ₂	$3[(R750 - R705) - 0.2(R750 - R550)(R750/R705)]$	[106]
Modified chlorophyll absorption ratio index	MCARI ₁	$[(R700 - R670) - 0.2(R700 - R550)](R700/R670)$	[15]
	MCARI ₂	$[(R750 - R705) - 0.2(R750 - R550)](R750/R705)$	[106]
	MCARI ₃	$[(R750 - R710) - 0.2(R750 - R550)](R750/R715)$	[106]
Optimized soil-adjusted vegetation index	OSAVI ₁	$(1 + 0.16)(R800 - R670)/(R800 + R670 + 0.16)$	[107]
	OSAVI ₂	$(1 + 0.16)(R750 - R705)/(R750 + R705 + 0.16)$	[106]
Combined TCARI/OSAVI	TCARI/OSAVI ₁	TCARI ₁ /OSAVI ₁	[105]
	TCARI/OSAVI ₂	TCARI ₂ /OSAVI ₂	[106]
Combined MCARI/OSAVI	MCARI/OSAVI ₁	MCARI ₁ /OSAVI ₁	[106]
	MCARI/OSAVI ₂	MCARI ₂ /OSAVI ₂	[106]
Triangular greenness index	TGI	$-0.5[190(R670-R550) - 120(R670 - R480)]$	[108]
Modified triangular vegetation index	MTVI	$1.2[1.2(R800 - R550) - 2.5(670 - R550)]$	[109]
MERIS terrestrial chlorophyll index	MTCI ₁	$(R750 - R710)/(R710 - R680)$	[110]
	MTCI ₂	$(R754 - R709)/(R709 - R681)$	
Spectral polygon vegetation index	SPVI	$0.4 \times [3.7(R800 - R670) - 1.2[R550 - R670]]$	[111]
Red edge position index	REP ₁	$700 + 45[(R670 + R780)/2 - R700]/(R740 - R700)$	[69]
	REP ₂	$700 + 40[(R670 + R780)/2 - R700]/(R740 - R700)$	[112]
-	VOG ₁	R740/R720	[113]
	VOG ₂	$(R734 - R747)/(R715 + R726)$	
	VOG ₃	$(R734 - R747)/(R715 + R720)$	
Optimal vegetation index	Vi _{opt}	$(1 + 0.45)(R800 + 1)/(R670 + 0.45)$	[114]

In supervised learning, feature selection is typically applied prior to the model development to reduce the data dimensionality, especially when the training set is small. The recursive feature

elimination (RFE) approach was a widely applied wrapper feature selection model and performed well in previous studies [46,115,116]. Therefore, it was adopted for feature selection in this study. It was performed iteratively by: (1) running an estimator to determine the initial feature importance scores; (2) removing the feature with the lowest importance score; and (3) assigning the ranking to the removed variable according to its removing order. This procedure was recursively repeated until the rankings were determined for all the input features. We implemented the entire process twice independently using two distinct estimators including SVR [117] and RF [118] to calculate the initial feature importance in step (1), and the final rankings were calculated based on the results from SVR-RFE and RF-RFE.

2.4. Ensemble Model Development

To enhance the prediction performance, an ensemble model was proposed based on a stacking strategy including the following two steps: (1) train and apply multiple machine learning models independently; and (2) combine multiple prediction results via a linear regressor [119]. It is important to select appropriate base learners to develop a successful ensemble model, and diversity is a critical condition, as the similarity between different models must be minimized for providing complementary information [60]. SVR, K-nearest neighbors (KNN) and RF are commonly used machine learning approaches with distinct principles, and their predictive ability for crop yield prediction have been assessed by many studies [120–122]. Therefore, we used these three models as the base learners and a brief description for each algorithm is presented below.

SVR is a supervised learning model characterized by the usage of kernels [123]. In SVR, the variables are firstly transformed from the original space to another space through a kernel function. Then, a linear function is determined to minimize errors between training data and the insensitive loss function. KNN regression assumes that similar samples exist in close proximity in the feature space. As an instance-based learning method, it estimates the response of an unknown sample by averaging the responses from K nearest neighbors in the training set. RF regression is a combination of multiple regression trees following the classical top-down procedure [124]. Each tree is generated using a bootstrap sample and learned in parallel and independent of each other [125,126]. The final estimation of RF is determined by averaging the predictions of all the independent trees.

In this study, we adopted a five-fold cross validation strategy [4,127,128] to create out-of-sample predictions. To test the robustness of all models, 50 repetitions of five-fold cross validation were performed, resulting in a total of 250 experiments (Figure 2). The R^2 , root mean square error (RMSE) and mean absolute deviation (MAE) were used to evaluate the model performance, and the equations of the indices are shown in Equations (1)–(3).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

where n is the number of samples, y_i and \hat{y}_i represent the observed and the predicted yields of sample i , \bar{y} and denotes the mean of observed yield. Models with higher R^2 and lower values of RMSE and MAE indicating better performance in prediction. In addition, a paired sample t-test, was adopted to evaluate the significance of the accuracy improvement of the ensemble model [129].

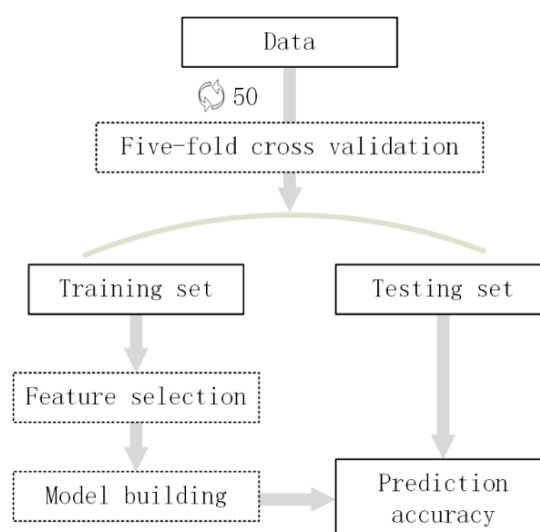


Figure 2. The experimental workflow of model development and validation.

3. Results

3.1. Yield Statistics and Spectral Profiles

Moisture contents were measured for each plot with an average of around 80% and a standard deviation (STD) of 2%, and the dry matter yield was used in this study. The yields varied at the two harvests, with an average of 2102.267 kg/ha in the August harvest, and a reduced average of 1042.149 kg/ha in September. The detailed yield statistics for each machinery treatment are shown in Table 3. Among different treatments, similar patterns were observed in both harvests. In general, compaction treatments with more traffic passes were associated with lower yields. For example, T1 and T7 had higher yields than others, and this is because T1 was simulating the pure mower operation using a single traffic pass, and T7 is the control group without any traffic compaction. In contrast, T6 had the lowest yield due to the most severe compaction. Moreover, earlier machinery treatments were found to have a less negative impact on the yield, by comparing the simulated silage (T2 and T3) and hay (T5 and T6) harvests.

Table 3. Yield statistics for seven machinery wheel traffic treatments.

Harvesting Time	Treatment	Mean (kg/ha)	Max. (kg/ha)	Min. (kg/ha)	STD (kg/ha)
August	T1	2256.319	3170.609	1333.134	450.226
	T2	2172.798	2495.764	1066.013	383.013
	T3	2074.944	2447.826	1282.724	324.696
	T4	2150.558	2752.013	923.186	500.141
	T5	2037.878	2729.032	1436.424	373.870
	T6	1808.317	2453.262	1095.171	357.561
	T7	2215.053	2686.530	1328.933	359.785
September	T1	1282.477	1528.347	1012.144	172.232
	T2	1077.874	1441.119	295.044	304.434
	T3	951.109	1262.709	510.520	230.796
	T4	1171.774	1475.466	386.226	284.171
	T5	868.823	1182.646	678.304	170.256
	T6	701.285	1049.704	256.495	228.078
	T7	1241.705	1439.142	562.412	228.325

To investigate the spectrum changes under compaction, the average alfalfa spectrum under each treatment (T1–T6) was compared with the control group (T7) which had no compaction applied, and the comparison results including the mean and standard deviation are shown in Figure 3. Similar to

other green vegetation, alfalfa plants generally have low reflectance in blue (450–500 nm) and red regions (600–700 nm) due to the chlorophyll absorption and high reflectance in NIR (800–900 nm) due to the scattering in spongy mesophyll [130]. In addition, we found that more compaction resulted in lower reflectance values in NIR due to the plant stress. For example, under the most severe wheel compaction, T6 showed the lowest average reflectance values within the NIR region (750–1000 nm), and the least spectra overlap with the control group T7.

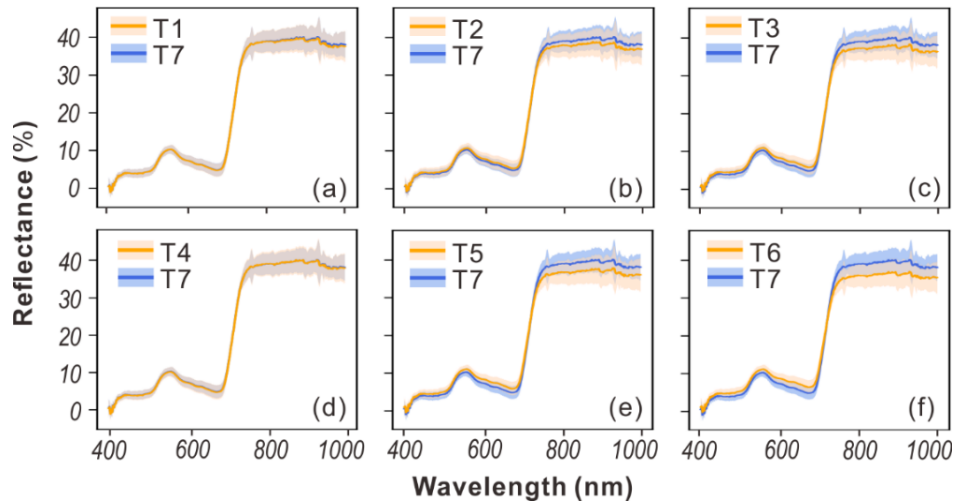


Figure 3. Comparison of average alfalfa spectrum with a standard deviation between each compaction treatment with the control group, (a) T1–T7, (b) T2–T7, (c) T3–T7, (d) T4–T7, (e) T5–T7 and (f) T6–T7.

3.2. Feature Importance

The 80 VIs were ranked using the RFE strategy described in Section 2.3. The ranking results from one experiment are shown in Table 4, and the ranking statistics over the 250 experiments are shown in Figure A1. Across the 250 experiments, the VIs ranked in the top and bottom were relatively stable. For example, Datt₁ and MCARI₁ mostly ranked highly, and they were originally developed for estimating the chlorophyll content of eucalyptus and corn leaves [15,86]. Two indices, MTCI₁ and MTCI₂, designed for a medium resolution imaging spectrometer [110] also had a stable and satisfactory performance. Out of all the 80 indices, there were 27 three-band or four-band VIs, and about 20 of them ranked in the top 40. This is likely because the appropriate combination of multiple bands could provide more spectral information. We also noticed that four integrated indices, including TCARI/OSAVI₁, TCARI/OSAVI₂, MCARI/OSAVI₁ and MCARI/OSAVI₂, ranked in the top 40, with TCARI/OSAVI₁ and MCARI/OSAVI₁ appearing in the top 20. This indicates that different types of VIs can be combined to complement each other for achieving enhanced performance. In addition, several narrow-band NDVIs, which have been widely adopted for estimating vegetation properties [68], were also effective in modeling the alfalfa yield.

3.3. Model Comparison and Performance

To further explore the high-performing features, we iteratively added the top one index into the machine learning model and updated the model training performance until all the 80 indices were included. The training accuracies obtained by the three base models (SVR, KNN and RF) were calculated, and the results of one experiment are shown in Figure 4. Among all the 250 experiments, RF performed the best, followed by KNN. As more features were included, the accuracies of all models improved at first and then remained stable after around 25 features were included. Therefore, we finally used the top 25 features for the ensemble model development.

Table 4. VI ranking using recursive feature elimination (RFE).

Feature	Ranking	Feature	Ranking
Datt1	1	SPVI	41
MCARI1	2	mSRI _[720,750]	42
MTCI2	3	VOG1	43
MCARI/OSAVI1	4	Carte2	44
MTCI1	5	TCARI1	45
REP2	6	MCARI2	46
PRI _[531,570]	7	Carte1	47
SR _[675,700]	8	NVI1	48
NDVI _[521,689]	9	NDVI _[471,584]	49
NDVI _[717,732]	10	NDVI _[667,740]	50
REP1	11	Datt2	51
TCARI/OSAVI1	12	mSR	52
NVI2	13	RDVI	53
TCARI2	14	SRI _[560,810]	54
TCARI/OSAVI2	15	NDVI _[710,750]	55
NDVI _[720,820]	16	SRI _[710,750]	56
Carte4	17	Datt3	57
NDVI _[734,750]	18	mND705	58
VOG3	19	mSRI _[710,780]	59
PRI _[528,567]	20	Gitelson1	60
VOG2	21	OSAVI1	61
NDRE	22	SRI _[705,750]	62
SRI _[533,565]	23	Gitelson2	63
EVI	24	NDVI _[717,770]	64
SRI _[720,735]	25	SRI _[670,800]	65
SRI _[629,734]	26	NDCI	66
DD	27	Carte3	67
MCARI/OSAVI2	28	SRI _[660,810]	68
CI	29	OSAVI2	69
SRI _[670,700]	30	mSRI _[550,780]	70
MTVI	31	NDVI _[705,750]	71
SRI _[700,750]	32	NDVI _[710,780]	72
NDVI _[550,760]	33	SRI _[550,750]	73
MCARI3	34	SRI _[706,755]	74
SRI _[717,741]	35	SRI _[550,760]	75
DCNI	36	SRI _[708,747]	76
TGI	37	mSR705	77
NDVI _[670,800]	38	SRI _[680,800]	78
SRI _[720,738]	39	GNDVI	79
Viopt	40	SRI _[690,752]	80

We trained all the four models (three base models, and one ensemble model) introduced in Section 2.4 using both the full and selected features on training samples, and evaluated the model performance on test samples. The test accuracies obtained from the 250 experiments are shown in Table 5. Satisfactory accuracies with an R^2 of more than 0.822 were achieved by all the approaches, demonstrating the effectiveness of these models in modeling alfalfa yield. The ensemble model outperformed all the base approaches, achieving an R^2 of 0.874 using the reduced features and an R^2 of 0.854 using the full features. Regarding the feature selection, the accuracies were improved for all the approaches. We also compared the results between using the selected VIs and the derivatives of full bands (Table A1). Again, the ensemble model outperformed other base models regardless of the features adopted, which further proved the effectiveness of the ensemble model proposed in this study. Besides, with the ensemble model, higher accuracies were achieved by using the selected VIs than using first or second derivatives of full bands. Moreover, we used a paired sample t-test to evaluate whether the methods are statistically different on the reported R^2 . Specifically, six statistical tests were

performed, and the results are shown in Table 6. The t-test was conducted between the ensemble model with each of the three base models (RF, SVR and KNN). The results showed that the accuracy improvement obtained by the ensemble model was statistically significant under both selected and full feature conditions.

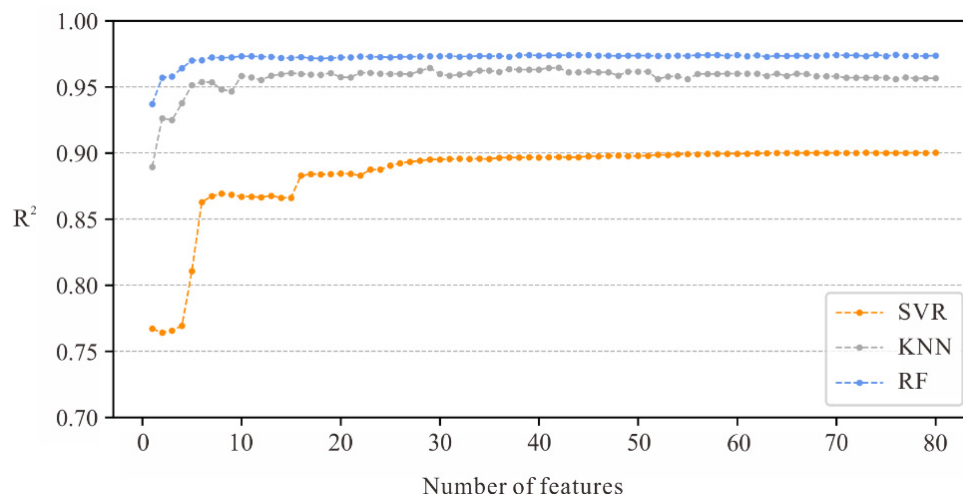


Figure 4. Model training accuracy as a function of the number of features.

Table 5. Test accuracies including mean and standard deviation of random forest (RF), support vector regression (SVR), K-nearest neighbors (KNN) and ensemble model in predicting the alfalfa yield.

Feature	Model	R ²	RMSE (kg/ha)	MAE (kg/ha)
Selected features	RF	0.833 (0.052)	252.912 (36.243)	185.317 (27.611)
	SVR	0.842 (0.042)	247.593 (36.269)	185.869 (27.128)
	KNN	0.850 (0.035)	241.430 (31.183)	183.557 (25.998)
	Ensemble	0.874 (0.034)	220.799 (32.169)	164.787 (24.673)
Full features	RF	0.822 (0.054)	261.552 (35.718)	191.602 (28.109)
	SVR	0.829 (0.042)	257.408 (30.260)	191.590 (23.569)
	KNN	0.822 (0.044)	262.907 (35.417)	198.293 (28.847)
	Ensemble	0.854 (0.036)	237.906 (32.152)	175.575 (25.300)

Table 6. Results from the paired sample t-test.

Feature	Model	t	p-Value
Selected features	Ensemble vs. RF	18.355	0.000
	Ensemble vs. SVR	16.890	0.000
	Ensemble vs. KNN	17.059	0.000
Full features	Ensemble vs. RF	15.935	0.000
	Ensemble vs. SVR	13.957	0.000
	Ensemble vs. KNN	20.255	0.000

The agreement between the observed and predicted yield for each model using the selected features is shown in Figure 5. Among the four models, the best agreement was found in the ensemble

model (Figure 5d), while the worst agreement was observed in RF (Figure 5a). Moreover, in Figure 5, two clusters could be identified and they respectively corresponded to the yields obtained from the two harvests. Again, comparing to the other approaches, the ensemble model demonstrated better agreement in both clusters.

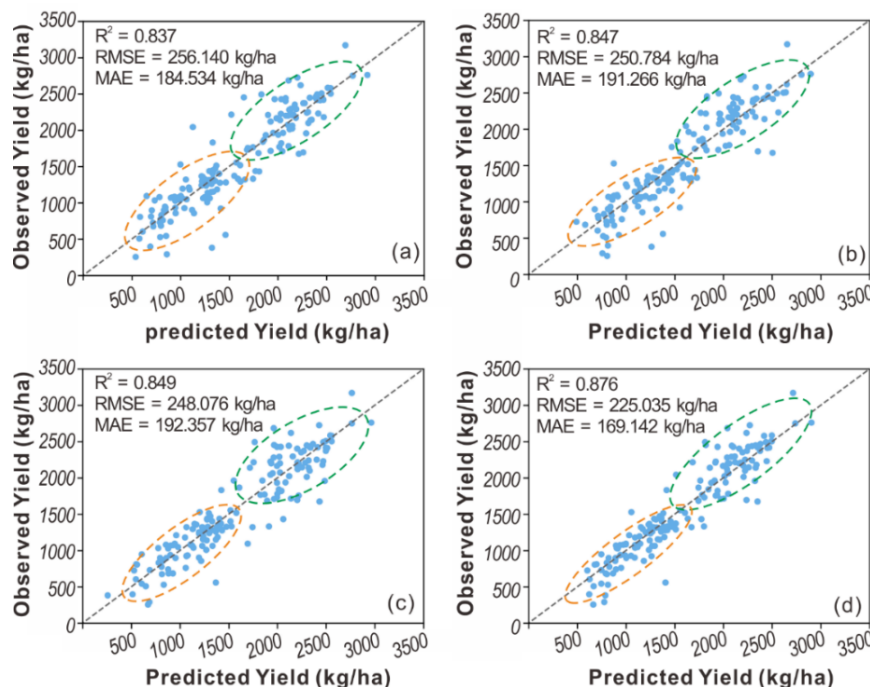


Figure 5. Scatter plots of observed vs. predicted yields from (a) RF, (b) SVR, (c) KNN and (d) ensemble model.

3.4. Model Adaptability for Different Compaction Treatments

We then further evaluated the model adaptability under different compaction treatments. To this end, we calculated the prediction accuracies for all seven treatments, and the results are shown in Table 7. In general, all models achieved good results and they performed particularly well under T1, T2, T3 and T5 treatments, with all the R^2 above 0.831. Additionally, we noticed that the ensemble model provided decent estimation accuracy under all the treatments, and it was superior to other base models in most cases. Although the accuracy varied across treatments, the ensemble model performed relatively stable, with an R^2 of more than 0.778 under each treatment, demonstrating its stronger adaptability across treatments. Moreover, it is worth mentioning that the ensemble performance can be directly affected by the base learners. For example, it achieved better performance in T3 and T5 when the base models showing higher accuracies.

Finally, the detailed yield modeling performance for each treatment was shown in the scatter plots in Figure 6. It is clear to see that T1 and T7 had higher yields than others, while T6 had the lowest yield due to the most severe machinery compaction. Among the four models, the ensemble model again showed the best agreement between the observed and the predicted yield under most treatments, and strong performance was exhibited under T3 and T5 treatments. Additionally, the figure also showed that both the high and low yields obtained from the two harvests could be well modeled using the ensemble approach.

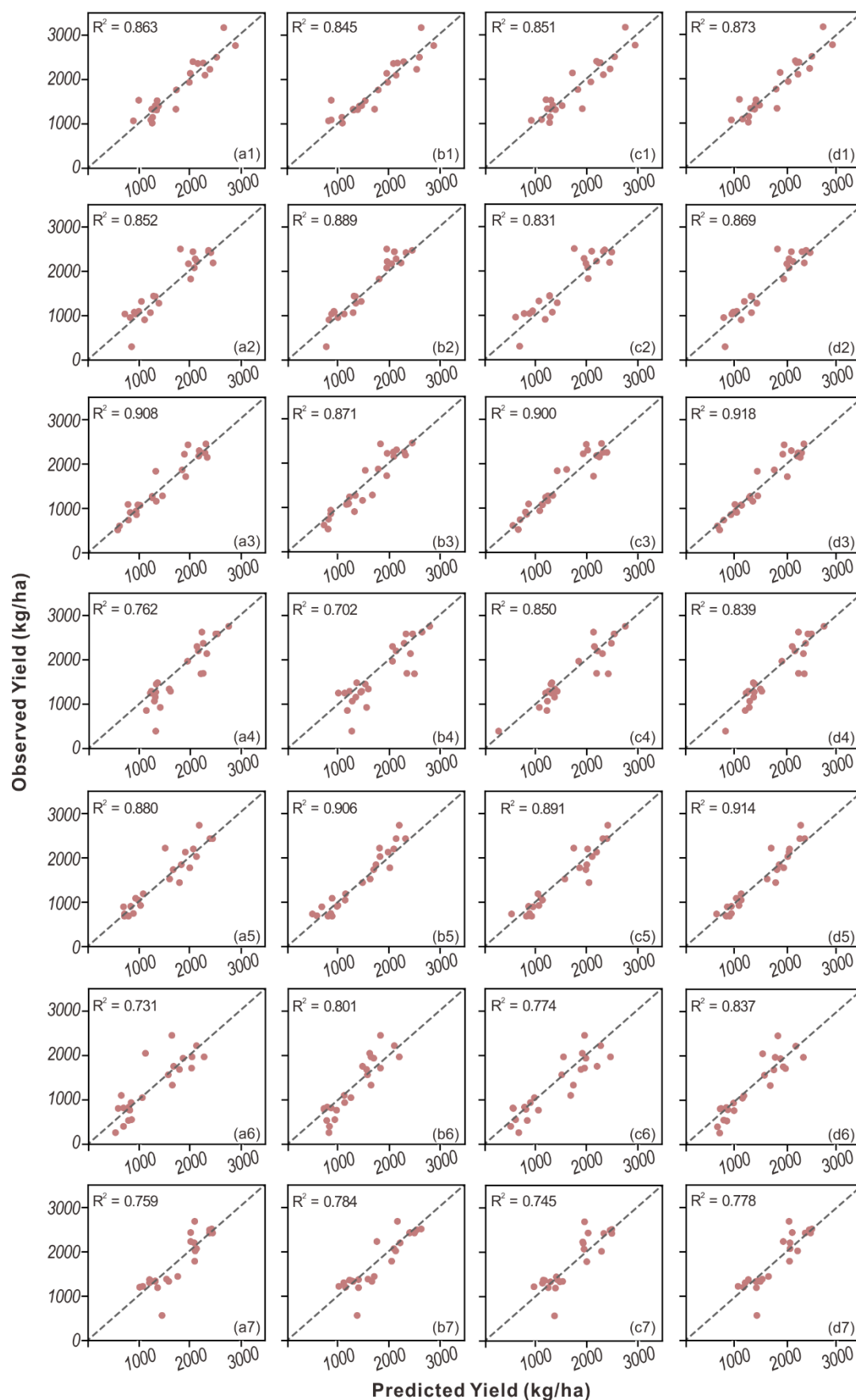


Figure 6. Scatter plots of observed vs. predicted yields from (a) RF, (b) SVR, (c) KNN and (d) ensemble model for groups with seven different compaction treatments: (1–7) T1–T7.

Table 7. Performance of RF, SVR, KNN and ensemble model in yield prediction for alfalfa under seven machinery compaction treatments.

Model	Metrics	T1	T2	T3	T4	T5	T6	T7
RF	R ²	0.863	0.852	0.908	0.762	0.880	0.731	0.759
	RMSE (kg/ha)	221.319	249.490	190.915	310.648	226.252	327.119	281.774
	MAE (kg/ha)	170.618	192.834	132.685	216.692	154.630	229.184	196.870
SVR	R ²	0.845	0.889	0.871	0.702	0.906	0.801	0.784
	RMSE (kg/ha)	235.242	215.894	225.637	347.210	200.501	281.140	266.777
	MAE (kg/ha)	166.508	168.490	172.771	254.350	162.607	228.031	185.659
KNN	R ²	0.851	0.831	0.900	0.850	0.891	0.774	0.745
	RMSE (kg/ha)	230.624	266.081	198.819	246.571	216.394	299.885	289.737
	MAE (kg/ha)	185.806	222.928	151.893	172.299	166.075	240.781	210.967
Ensemble	R ²	0.873	0.869	0.918	0.839	0.914	0.837	0.778
	RMSE (kg/ha)	212.574	234.157	180.469	255.500	192.307	254.738	270.159
	MAE (kg/ha)	159.047	185.788	126.512	191.685	142.453	191.387	189.763

4. Discussion

4.1. Selection of the Vegetation Indices

In this study, we explored 80 hyperspectral narrow-band indices, and approximately 95% of these indices used the red-edge wavelengths in their calculations. To reduce the data dimensionality, all the 80 VIs were ranked based on the REF approach, and several representative indices were discussed below.

First, we noticed that three-band or four-band indices ranked importantly and about six of them were among the top ten indices. This is likely because the multi-band indices tend to provide more spectral information and improve the robustness and precision in assessing plant traits [92]. For example, Datt₁, MCARI₁, MTCI₁ and MTCI₂ are four three-band indices, and they were relatively stable in the top ten among the 250 experiments. Their strong performance is mainly due to their sensitivity to the chlorophyll content which is strongly correlated with the photosynthetic capacity of the plants [131]. In previous studies, integrated indices such as MCARI/OSAVI and TCARI/OSAVI were found to outperform TCARI and MCARI, and this is mainly because combining the MCARI or TCARI with the optimized SAVI (OSAVI) can help reduce the background effects [15]. However, this pattern was not observed in this study as the background pixels (e.g., soil) were already excluded in the data pre-processing step. Another multi-band index REP is calculated using four narrow bands to ascertain the position of maximum slope in the red-NIR region, and its potential of estimating biomass has been identified in previous studies [76,132,133]. Two REP indices (REP₁ and REP₂) were considered in this study, and both were selected for the model development.

In addition, various narrow-band NDVIs have been evaluated by researchers for estimating different plant traits [68,88,111,134] and several of them were also demonstrated to be effective in modeling alfalfa yield in this study. Another common index is the physiological reflectance index (PRI), which is typically used for indicating the photosynthetic efficiency [78]. In this study, PRI_[531,570] also ranked importantly and was selected for the model development. Besides, we also studied 19 SRIs with multiple band combinations, and only about three of them were selected as predictors, and none of them ranked in the top ten. This is likely because the information provided by SRIs are already included in other VIs with more spectral bands or complex mathematical structures.

4.2. Advantages of the Ensemble Model

Instead of using a single machine learning model, we developed an ensemble model by combining three base learners in this study. The results showed that the ensemble model outperformed the other approaches significantly when using both all the VIs and the selected ones. With the selected features, the ensemble method obtained 0.874 in R², 220.799 kg/ha in RMSE and 164.787 kg/ha in MSE, achieving an increase of 2.8%, 8.5%, and 10.2%, respectively, in comparison with the KNN which was the best performing single model. In addition, we further evaluated the model performance on seven

compaction treatments. Again, the ensemble model yielded more adaptive and satisfactory results than base models with R^2 exceeding 0.778 on all the subsets.

As demonstrated previously, in ensemble learning, sufficiency and diversity are the two main principles in selecting the base models [135]. It means that each base learner should possess good predicting capabilities, but at the same time, the dependence between the models must be minimized for providing complementary information [60]. This is reasonable because the ensemble method combines the individual predictions, and therefore the performance of each base model can affect the final fusion results. On the other hand, limited additional information would be gained by combining the high-performance models which are similar. Based on the two conditions, we employed SVR, KNN and RF which have completely different training mechanisms, as the base learners in this study, and their parameters were optimized strategically for achieving the best performance. The experimental results further demonstrate the effectiveness of the base model selection.

4.3. Effects of Machinery Compaction

Previous studies showed that alfalfa is one of the most susceptible forages to machinery traffic which may lead to a reduction in the next cutting yield. In this study, seven levels of compaction were explored, and two main insights were found from the experiments. First, we noticed that the yield loss is positively correlated with the severity of the compaction. The more traffic passes were applied, the greater the observed yield loss. This is mainly because the soil compaction caused by the wheel traffic can lead to reduced macropore air permeability and water infiltration, and thus negatively affect the plant root development and lower the yield [12]. Second, the later application of machinery traffic practices tends to increase the yield loss, which could be noticed by comparing the yields from hay and silage harvests. This is because delayed wheel traffic is more likely to cause physical damage to regrowth shoot and therefore decrease the yield [9]. Similar findings were also noticed in other studies [9–11,135,136]. For example, wheel traffic was found to be a contributing factor to alfalfa yield loss in a field study conducted in Nevada, and fewer and earlier machinery treatments were recommended to reduce the damage [10]. A field experiment carried out in Auchincruive, Scotland, also demonstrated that both frequent and delayed wheel passes caused reductions in herbage yield [135].

5. Conclusions

Alfalfa is an important forage crop in the U.S., and it plays an important role in the food supply chain as feedstock for animals. Pre-harvest insight to yield can help optimize management practices. In this study, we developed an ensemble-based machine learning model for alfalfa yield prediction using UAV-based hyperspectral imagery. The narrow-band hyperspectral indices were extracted and the most important ones were selected for the model development. The results showed that the ensemble model outperformed other base models, and the highest accuracy was achieved when using the reduced features. In addition, we examined the model performance under different compaction treatments, and again, the best performance was achieved by the ensemble model, indicating its stronger adaptability in comparison with other approaches. Moreover, we also found that the yield loss is positively correlated with the compaction severity, while earlier compaction after harvest can help reduce the impact. Our study demonstrated the efficacy of using hyperspectral images for modeling alfalfa yield, and for future work, we will incorporate environmental factors, such as climate variables and soil properties, into the modeling process and to further enhance the prediction performance.

Author Contributions: Conceptualization, L.F., Z.Z., J.D. and B.L.; Data Collection, Y.M., Z.Z., P.W., J.D. and B.L.; Methodology, L.F. and Z.Z.; Validation, L.F., Y.M., and Z.Z.; Writing—original draft, L.F. and Z.Z.; Writing—review and editing, L.F., Z.Z., Y.M. and Q.D., P.W., J.D. and B.L. All authors have read and agreed to the published version of the manuscript.

Funding: Support for this research was provided by the National Institute of Food and Agriculture, United States Department of Agriculture, under ID number WIS03026.

Acknowledgments: This work was partly supported by the China Scholarship Council (NO.201906270096).

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

Appendix A. Appendix

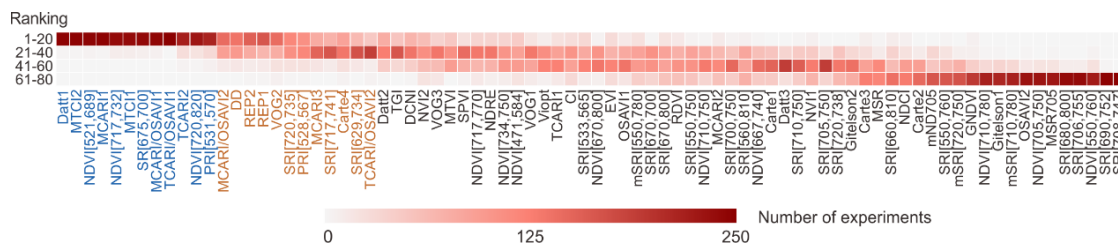


Figure A1. Statistics of VI rankings in 250 experiments.

Table A1. Test accuracies of RF, SVR, KNN and ensemble model in predicting the alfalfa yield.

Feature	Model	R ²	RMSE (kg/ha)	MAE (kg/ha)
First derivatives of full bands	RF	0.848	249.623	187.215
	SVR	0.845	251.307	198.938
	KNN	0.823	267.859	210.981
	Ensemble	0.869	231.887	180.150
Second derivatives of full bands	RF	0.820	270.324	209.001
	SVR	0.814	275.077	217.356
	KNN	0.800	283.308	217.900
	Ensemble	0.836	258.841	201.951
First and second derivatives of full bands	RF	0.856	241.011	181.244
	SVR	0.854	243.162	195.826
	KNN	0.803	281.108	217.741
	Ensemble	0.874	226.434	177.255
Selected VIs	RF	0.833	252.912	185.317
	SVR	0.842	247.593	185.869
	KNN	0.850	241.430	183.557
	Ensemble	0.874	220.799	164.787

References

- Radović, J.; Sokolović, D.; Marković, J. Alfalfa-most important perennial forage legume in animal husbandry. *Biotechnol. Anim. Husb.* **2009**, *25*, 465–475. [CrossRef]
- United States Department of Agriculture National Agricultural Statistics Service. Available online: <https://www.nass.usda.gov/> (accessed on 3 March 2020).
- Andrzejewska, J.; Contreras-Govea, F.E.; Albrecht, K.A. Field prediction of alfalfa (*Medicago sativa* L.) fibre constituents in northern Europe. *Grass Forage Sci.* **2014**, *69*, 348–355. [CrossRef]
- Noland, R.L.; Wells, M.S.; Coulter, J.A.; Tiede, T.; Baker, J.M.; Martinson, K.L.; Sheaffer, C.C. Estimating alfalfa yield and nutritive value using remote sensing and air temperature. *Field Crops Res.* **2018**, *222*, 189–196. [CrossRef]
- Summers, C.G.; Putnam, D.H. *Irrigated Alfalfa Management for Mediterranean and Desert Zones*; UCANR Publications: Davis, CA, USA, 2008; Volume 3512.
- Głab, T. Effect of Soil Compaction on Root System Morphology and Productivity of Alfalfa (*Medicago sativa* L.). *Pol. J. Environ. Stud.* **2011**, *20*, 1473–1480.
- Harris, P.A.; Ellis, A.D.; Fradinho, M.J.; Jansson, A.; Julliand, V.; Luthersson, N.; Santos, A.S.; Vervuert, I. Feeding conserved forage to horses: Recent advances and recommendations. *Animal* **2017**, *11*, 958–967. [CrossRef]

8. Manyawu, G. *Principles of Silage Making*; International Livestock Research Institute (ILRI): Nairobi, Kenya, 2016.
9. Undersander, D.; Cosgrove, D. *Alfalfa Management Guide*; American Society of Agronomy Crop Science Society of America Soil Science: Madison, WI, USA, 2011.
10. Schmierer, J.; Putnam, D.; Undersander, D.; Liu, J.; Meister, H. Wheel Traffic in Alfalfa—What do We Know? What Can We Do About It? In Proceedings of the National Alfalfa Symposium, San Diego, CA, USA, 13–15 December 2004; pp. 13–15.
11. Rechel, E.A.; Meek, B.D.; DeTar, W.R.; Carter, L.M. Alfalfa yield as affected by harvest traffic and soil compaction in a sandy loam soil. *J. Prod. Agric.* **1991**, *4*, 241–246. [[CrossRef](#)]
12. Undersander, D.; Moutray, J. Effect of Wheel Traffic on Alfalfa Yield. Available online: <https://fyi.extension.wisc.edu/forage/effect-of-wheel-traffic-on-alfalfa/> (accessed on 3 March 2020).
13. Orloff, S.; Putnam, D. Adjusting alfalfa cutting schedules for economic conditions. In Proceedings of the 40th California Alfalfa & Forage and Corn/Cereal Silage Mini-Symposium, Visalia, CA, USA, 1 December 2010; pp. 1–2.
14. Undersander, D. Alfalfa Yield and Stand. Available online: <https://fyi.extension.wisc.edu/forage/alfalfa-yield-and-stand/> (accessed on 3 March 2020).
15. Daughtry, C.S.T.; Walthall, C.L.; Kim, M.S.; De Colstoun, E.B.; McMurtrey Iii, J.E. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [[CrossRef](#)]
16. Geipel, J.; Link, J.; Claupein, W. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. *Remote Sens.* **2014**, *6*, 10335–10355. [[CrossRef](#)]
17. Lauer, J. Methods for Calculating Corn Yield. 2002, Volume 33. Available online: <http://corn.agronomy.wisc.edu/AA/A033.aspx> (accessed on 3 March 2020).
18. Ma, Y.; Kang, Y.; Ozdogan, M.; Zhang, Z. County-level corn yield prediction using deep transfer learning. *AGUFM* **2019**, *2019*, B54D-02.
19. Wang, Y.; Zhang, Z.; Feng, L.; Du, Q.; Runge, T. Combining Multi-Source Data and Machine Learning Approaches to Predict Winter Wheat Yield in the Conterminous United States. *Remote Sens.* **2020**, *12*, 1232. [[CrossRef](#)]
20. Pan, G.; Sun, G.-J.; Li, F.-M. Using QuickBird imagery and a production efficiency model to improve crop yield estimation in the semi-arid hilly Loess Plateau, China. *Environ. Model. Softw.* **2009**, *24*, 510–516. [[CrossRef](#)]
21. Wahab, I.; Hall, O.; Jirstrom, M. Remote Sensing of Yields: Application of UAV Imagery-Derived NDVI for Estimating Maize Vigor and Yields in Complex Farming Systems in Sub-Saharan Africa. *Drones* **2018**, *2*, 28. [[CrossRef](#)]
22. Su, W.; Zhang, M.; Bian, D.; Liu, Z.; Huang, J.; Wang, W.; Wu, J.; Guo, H. Phenotyping of Corn Plants Using Unmanned Aerial Vehicle (UAV) Images. *Remote Sens.* **2019**, *11*, 2021. [[CrossRef](#)]
23. Kayad, A.G.; Al-Gaadi, K.A.; Tola, E.; Madugundu, R.; Zeyada, A.M.; Kalaitzidis, C. Assessing the spatial variability of alfalfa yield using satellite imagery and ground-based data. *PLoS ONE* **2016**, *11*, e0157166. [[CrossRef](#)]
24. Sanches, G.M.; Duft, D.G.; Kölln, O.T.; Luciano, A.C.d.S.; De Castro, S.G.Q.; Okuno, F.M.; Franco, H.C.J. The potential for RGB images obtained using unmanned aerial vehicle to assess and predict yield in sugarcane fields. *Int. J. Remote Sens.* **2018**, *39*, 5402–5414. [[CrossRef](#)]
25. Yun, H.S.; Park, S.H.; Kim, H.-J.; Lee, W.D.; Lee, K.D.; Hong, S.Y.; Jung, G.H. Use of unmanned aerial vehicle for multi-temporal monitoring of soybean vegetation fraction. *J. Biosyst. Eng.* **2016**, *41*, 126–137. [[CrossRef](#)]
26. Zhou, X.; Zheng, H.B.; Xu, X.Q.; He, J.Y.; Ge, X.K.; Yao, X.; Cheng, T.; Zhu, Y.; Cao, W.X.; Tian, Y.C. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 246–255. [[CrossRef](#)]
27. Hassan, M.; Yang, M.; Rasheed, A.; Jin, X.; Xia, X.; Xiao, Y.; He, Z. Time-series multispectral indices from unmanned aerial vehicle imagery reveal senescence rate in bread wheat. *Remote Sens.* **2018**, *10*, 809. [[CrossRef](#)]

28. Romero, M.; Luo, Y.; Su, B.; Fuentes, S. Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. *Comput. Electron. Agric.* **2018**, *147*, 109–117. [\[CrossRef\]](#)
29. Maresma, Á.; Ariza, M.; Martínez, E.; Lloveras, J.; Martínez-Casasnovas, J. Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAV service. *Remote Sens.* **2016**, *8*, 973. [\[CrossRef\]](#)
30. Cárdenas, D.A.G.; Valencia, J.A.R.; Velásquez, D.F.A.; Gonzalez, J.R.P. Dynamics of the Indices NDVI and GNDVI in a Rice Growing in Its Reproduction Phase from Multi-spectral Aerial Images Taken by Drones. In *Advances in Intelligent Systems and Computing, Proceedings of the 2nd International Conference of ICT for Adapting Agriculture to Climate Change (AACC'18), Cali, Colombia, 21–23 November 2018*; Springer: Cham, Switzerland, 2018; pp. 106–119.
31. Shen, X.; Cao, L.; Yang, B.; Xu, Z.; Wang, G. Estimation of forest structural attributes using spectral indices and point clouds from UAS-based multispectral and RGB imageries. *Remote Sens.* **2019**, *11*, 800. [\[CrossRef\]](#)
32. Nidamanuri, R.R.; Zbell, B. Use of field reflectance data for crop mapping using airborne hyperspectral image. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 683–691. [\[CrossRef\]](#)
33. Yang, C.; Everitt, J.H.; Bradford, J.M. Airborne hyperspectral imagery and linear spectral unmixing for mapping variation in crop yield. *Precis. Agric.* **2007**, *8*, 279–296. [\[CrossRef\]](#)
34. Mewes, T.; Franke, J.; Menz, G. Spectral requirements on airborne hyperspectral remote sensing data for wheat disease detection. *Precis. Agric.* **2011**, *12*, 795. [\[CrossRef\]](#)
35. Ray, S.S.; Jain, N.; Arora, R.K.; Chavan, S.; Panigrahy, S. Utility of hyperspectral data for potato late blight disease detection. *J. Indian Soc. Remote Sens.* **2011**, *39*, 161. [\[CrossRef\]](#)
36. Kim, Y.; Glenn, D.M.; Park, J.; Ngugi, H.K.; Lehman, B.L. Hyperspectral image analysis for water stress detection of apple trees. *Comput. Electron. Agric.* **2011**, *77*, 155–160. [\[CrossRef\]](#)
37. Ranjan, R.; Chopra, U.K.; Sahoo, R.N.; Singh, A.K.; Pradhan, S. Assessment of plant nitrogen stress in wheat (*Triticum aestivum* L.) through hyperspectral indices. *Int. J. Remote Sens.* **2012**, *33*, 6342–6360. [\[CrossRef\]](#)
38. Jin, X.; Kumar, L.; Li, Z.; Xu, X.; Yang, G.; Wang, J. Estimation of winter wheat biomass and yield by combining the AquaCrop model and field hyperspectral data. *Remote Sens.* **2016**, *8*, 972. [\[CrossRef\]](#)
39. Montesinos-López, O.A.; Montesinos-López, A.; Crossa, J.; de los Campos, G.; Alvarado, G.; Suchismita, M.; Rutkoski, J.; González-Pérez, L.; Burgueño, J. Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods* **2017**, *13*, 4.
40. Aguater, F.M.; Trachsel, S.; Pérez, L.G.; Burgueño, J.; Crossa, J.; Balzarini, M.; Gouache, D.; Bogard, M.; Campos, G.D.L. Use of hyperspectral image data outperforms vegetation indices in prediction of maize yield. *Crop Sci.* **2017**, *57*, 2517–2524. [\[CrossRef\]](#)
41. Kawamura, K.; Ikeura, H.; Phongchanmaixay, S.; Khanthavong, P. Canopy hyperspectral sensing of paddy fields at the booting stage and PLS regression can assess grain yield. *Remote Sens.* **2018**, *10*, 1249. [\[CrossRef\]](#)
42. Yang, C.; Everitt, J.H.; Bradford, J.M.; Murden, D. Comparison of airborne multispectral and hyperspectral imagery for estimating grain sorghum yield. *Trans. ASABE* **2009**, *52*, 641–649. [\[CrossRef\]](#)
43. Oehlschläger, J.; Schmidhalter, U.; Noack, P.O. UAV-Based Hyperspectral Sensing for Yield Prediction in Winter Barley. In *Proceedings of the 2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Amsterdam, The Netherlands, 23–26 September 2018*; pp. 1–4.
44. Kanning, M.; Kühling, I.; Trautz, D.; Jarmer, T. High-resolution UAV-based hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction. *Remote Sens.* **2018**, *10*, 2000. [\[CrossRef\]](#)
45. Moghimi, A.; Yang, C.; Anderson, J.A. Aerial hyperspectral imagery and deep neural networks for high-throughput yield phenotyping in wheat. *arXiv* **2019**, arXiv:1906.09666.
46. Zhao, J.; Karimzadeh, M.; Masjedi, A.; Wang, T.; Zhang, X.; Crawford, M.M.; Ebert, D.S. FeatureExplorer: Interactive Feature Selection and Exploration of Regression Models for Hyperspectral Images. In *Proceedings of the 2019 IEEE Visualization Conference (VIS), Vancouver, BC, Canada, 20–25 October 2019*; pp. 161–165.
47. Schreiber, M.M.; Miles, G.E.; Holt, D.A.; Bula, R.J. Sensitivity Analysis of SIMED 1. *Agron. J.* **1978**, *70*, 105–108. [\[CrossRef\]](#)
48. Fick, G.W. *ALSIM 1 (Level 2) User's Manual*; Department of Agronomy, Cornell University: Ithaca, NY, USA, 1981.
49. Denison, R.F.; Loomis, R.S. *An Integrative Physiological Model of Alfalfa Growth and Development*; Publication/University of California, Division of Agriculture and Natural Resources (USA): Davis, CA, USA, 1989.

50. Bourgeois, G.; Savoie, P.; Girard, J.-M. Evaluation of an alfalfa growth simulation model under Quebec conditions. *Agric. Syst.* **1990**, *32*, 1–12. [\[CrossRef\]](#)
51. Malik, W.; Boote, K.J.; Hoogenboom, G.; Caverso, J.; Dechmi, F. Adapting the CROPGRO model to simulate alfalfa growth and yield. *Agron. J.* **2018**, *110*, 1777–1790. [\[CrossRef\]](#)
52. Cai, Y.; Moore, K.; Pellegrini, A.; Elhaddad, A.; Lessel, J.; Townsend, C.; Solak, H.; Semret, N. Crop yield predictions-high resolution statistical model for intra-season forecasts applied to corn in the US. In Proceedings of the American Geophysical Union 2017 Fall Meeting, New Orleans, LA, USA, 13 December 2017.
53. Zhang, Z.; Jin, Y.; Chen, B.; Brown, P. California Almond Yield Prediction at the Orchard Level With a Machine Learning Approach. *Front. Plant Sci.* **2019**, *10*, 809. [\[CrossRef\]](#)
54. Michel, L.; Makowski, D. Comparison of statistical models for analyzing wheat yield time series. *PLoS ONE* **2013**, *8*, e78615. [\[CrossRef\]](#)
55. Gandhi, N.; Armstrong, L.J.; Petkar, O.; Tripathy, A.K. Rice crop yield prediction in India using support vector machines. In Proceedings of the 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), Khon Kaen, Thailand, 13–15 July 2016; pp. 1–5.
56. Ali, I.; Cawkwell, F.; Green, S.; Dwyer, N. Application of statistical and machine learning models for grassland yield estimation based on a hypertemporal satellite remote sensing time series. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 5060–5063.
57. Pal, M. Ensemble learning with decision tree for remote sensing classification. *World Acad. Sci. Eng. Technol.* **2007**, *36*, 258–260.
58. Zhang, Z.; Pasolli, E.; Crawford, M.M.; Tilton, J.C. An active learning framework for hyperspectral image classification using hierarchical segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *9*, 640–654. [\[CrossRef\]](#)
59. Zhang, Z.; Pasolli, E.; Crawford, M.M. An Adaptive Multiview Active Learning Approach for Spectral-Spatial Classification of Hyperspectral Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 2557–2570. [\[CrossRef\]](#)
60. Zhou, Z.-H. Ensemble Learning. *Encycl. Biom.* **2009**, *1*, 270–273.
61. Aghighi, H.; Azadbakht, M.; Ashourloo, D.; Shahrabadi, H.S.; Radiom, S. Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 4563–4577. [\[CrossRef\]](#)
62. Feng, P.; Ma, J.; Sun, C.; Xu, X.; Ma, Y. A novel dynamic Android malware detection system with ensemble learning. *IEEE Access* **2018**, *6*, 30996–31011. [\[CrossRef\]](#)
63. Ju, C.; Bibaut, A.; van der Laan, M. The relative performance of ensemble methods with deep convolutional neural networks for image classification. *J. Appl. Stat.* **2018**, *45*, 2800–2818. [\[CrossRef\]](#)
64. U.S. Climate Data. Available online: <https://www.usclimatedata.com/#> (accessed on 3 March 2020).
65. Habib, A.; Zhou, T.; Masjedi, A.; Zhang, Z.; Flatt, J.E.; Crawford, M. Bore-sight calibration of GNSS/INS-assisted push-broom hyperspectral scanners on UAV platforms. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1734–1749. [\[CrossRef\]](#)
66. Peñuelas, J.; Filella, I. Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends Plant Sci.* **1998**, *3*, 151–156. [\[CrossRef\]](#)
67. Thompson, W.M.O. *The Whitefly, Bemisia Tabaci (Homoptera: Aleyrodidae) Interaction with Geminivirus-Infected Host Plants: Bemisia Tabaci, Host Plants and Geminiviruses*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011.
68. Liang, L.; Di, L.; Zhang, L.; Deng, M.; Qin, Z.; Zhao, S.; Lin, H. Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. *Remote Sens. Environ.* **2015**, *165*, 123–134. [\[CrossRef\]](#)
69. Van Der Meij, B.; Kooistra, L.; Suomalainen, J.; Barel, J.M.; De Deyn, G.B. Remote sensing of plant trait responses to field-based plant-soil feedback using UAV-based optical sensors. *Biogeosciences* **2017**, *14*, 733. [\[CrossRef\]](#)
70. Zhao, B.; Ma, B.L.; Hu, Y.; Liu, J. Characterization of nitrogen and water status in oat leaves using optical sensing approach. *J. Sci. Food Agric.* **2015**, *95*, 367–378. [\[CrossRef\]](#)
71. Yu, K.; Gnyp, M.L.; Gao, J.; Miao, Y.; Chen, X.; Bareth, G. Using Partial Least Squares (PLS) to Estimate Canopy Nitrogen and Biomass of Paddy Rice in China's Sanjiang Plain. In Proceedings of the Workshop on UAV-based Remote Sensing Methods for Monitoring Vegetation, Cologne, Germany, 14 April 2014; pp. 99–103.

72. Tucker, C.J. *Red and Photographic Infrared Linear Combinations for Monitoring Vegetation*; NASA Goddard Space Flight Center: Greenbelt, MD, USA, 1978.
73. Gitelson, A.; Merzlyak, M.N. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B Biol.* **1994**, *22*, 247–252. [[CrossRef](#)]
74. Wu, C.; Niu, Z.; Tang, Q.; Huang, W.; Rivard, B.; Feng, J. Remote estimation of gross primary production in wheat using chlorophyll-related vegetation indices. *Agric. For. Meteorol.* **2009**, *149*, 1015–1021. [[CrossRef](#)]
75. Kooistra, L.; Suomalainen, J.; Iqbal, S.; Franke, J.; Wenting, P.; Bartholomeus, H.; Mùcher, S.; Becker, R. Crop monitoring using a light-weight hyperspectral mapping system for unmanned aerial vehicles: First results for the 2013 season. In Proceedings of the Proceedings of 2013 Workshop on UAV-Based Remote Sensing Methods for Monitoring Vegetation, Cologne, Germany, 9–10 September 2013; p. 5158.
76. Mutanga, O.; Skidmore, A.K. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* **2004**, *25*, 3999–4014. [[CrossRef](#)]
77. Thenkabail, P.S.; Smith, R.B.; De Pauw, E. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sens. Environ.* **2000**, *71*, 158–182. [[CrossRef](#)]
78. Gamon, J.A.; Penuelas, J.; Field, C.B. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [[CrossRef](#)]
79. Penuelas, J.; Baret, F.; Filella, I. Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica* **1995**, *31*, 221–230.
80. Rodriguez, D.; Fitzgerald, G.J.; Belford, R.; Christensen, L.K. Detection of nitrogen deficiency in wheat from spectral reflectance indices and basic crop eco-physiological concepts. *Aust. J. Agric. Res.* **2006**, *57*, 781–789. [[CrossRef](#)]
81. 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](#)]
82. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [[CrossRef](#)]
83. Roujean, J.-L.; Breon, F.-M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [[CrossRef](#)]
84. Marshak, A.; Knyazikhin, Y.; Davis, A.B.; Wiscombe, W.J.; Pilewskie, P. Cloud-vegetation interaction: Use of normalized difference cloud index for estimation of cloud optical thickness. *Geophys. Res. Lett.* **2000**, *27*, 1695–1698. [[CrossRef](#)]
85. Zarco-Tejada, P.J.; Pushnik, J.C.; Dobrowski, S.; Ustin, S.L. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sens. Environ.* **2003**, *84*, 283–294. [[CrossRef](#)]
86. Datt, B. A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using Eucalyptus leaves. *J. Plant Physiol.* **1999**, *154*, 30–36. [[CrossRef](#)]
87. Le Maire, G.; Francois, C.; Dufrene, E. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sens. Environ.* **2004**, *89*, 1–28. [[CrossRef](#)]
88. Chen, P.; Haboudane, D.; Tremblay, N.; Wang, J.; Vigneault, P.; Li, B. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sens. Environ.* **2010**, *114*, 1987–1997. [[CrossRef](#)]
89. Gitelson, A.A.; Buschmann, C.; Lichtenthaler, H.K. The chlorophyll fluorescence ratio F735/F700 as an accurate measure of the chlorophyll content in plants. *Remote Sens. Environ.* **1999**, *69*, 296–302. [[CrossRef](#)]
90. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [[CrossRef](#)]
91. Carter, G.A. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Remote Sens.* **1994**, *15*, 697–703. [[CrossRef](#)]
92. Tian, Y.C.; Yao, X.; Yang, J.; Cao, W.X.; Hannaway, D.B.; Zhu, Y. Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground-and space-based hyperspectral reflectance. *Field Crops Res.* **2011**, *120*, 299–310. [[CrossRef](#)]
93. Gitelson, A.A.; Merzlyak, M.N. Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* **1997**, *18*, 2691–2697. [[CrossRef](#)]

94. Xue, L.; Cao, W.; Luo, W.; Dai, T.; Zhu, Y. Monitoring leaf nitrogen status in rice with canopy spectral reflectance. *Agron. J.* **2004**, *96*, 135–142. [\[CrossRef\]](#)
95. Zhu, Y.; Yao, X.; Tian, Y.; Liu, X.; Cao, W. Analysis of common canopy vegetation indices for indicating leaf nitrogen accumulations in wheat and rice. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 1–10. [\[CrossRef\]](#)
96. McMurtrey Iii, J.E.; Chappelle, E.W.; Kim, M.S.; Meisinger, J.J.; Corp, L.A. Distinguishing nitrogen fertilization levels in field corn (*Zea mays* L.) with actively induced fluorescence and passive reflectance measurements. *Remote Sens. Environ.* **1994**, *47*, 36–44. [\[CrossRef\]](#)
97. Chappelle, E.W.; Kim, M.S.; McMurtrey Iii, J.E. Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll a, chlorophyll b, and carotenoids in soybean leaves. *Remote Sens. Environ.* **1992**, *39*, 239–247. [\[CrossRef\]](#)
98. Gupta, R.K.; Vijayan, D.; Prasad, T.S. Comparative analysis of red-edge hyperspectral indices. *Adv. Space Res.* **2003**, *32*, 2217–2222. [\[CrossRef\]](#)
99. Zarco-Tejada, P.J.; Miller, J.R. Land cover mapping at BOREAS using red edge spectral parameters from CASI imagery. *J. Geophys. Res. Atmos.* **1999**, *104*, 27921–27933. [\[CrossRef\]](#)
100. Gitelson, A.A.; Vina, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*. [\[CrossRef\]](#)
101. Gitelson, A.A.; Viña, A.; Arkebauer, T.J.; Rundquist, D.C.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **2003**, *30*. [\[CrossRef\]](#)
102. Chen, J.M. Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Can. J. Remote Sens.* **1996**, *22*, 229–242. [\[CrossRef\]](#)
103. Gupta, R.K.; Vijayan, D.; Prasad, T.S. New hyperspectral vegetation characterization parameters. *Adv. Space Res.* **2001**, *28*, 201–206. [\[CrossRef\]](#)
104. Huete, A.; Justice, C.; Liu, H. Development of vegetation and soil indices for MODIS-EOS. *Remote Sens. Environ.* **1994**, *49*, 224–234. [\[CrossRef\]](#)
105. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [\[CrossRef\]](#)
106. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [\[CrossRef\]](#)
107. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [\[CrossRef\]](#)
108. Hunt, E.R., Jr.; Doraiswamy, P.C.; McMurtrey, J.E.; Daughtry, C.S.T.; Perry, E.M.; Akhmedov, B. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 103–112. [\[CrossRef\]](#)
109. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [\[CrossRef\]](#)
110. Dash, J.; Curran, P.J. The MERIS terrestrial chlorophyll index. *Int. J. Remote Sens.* **2004**, *25*, 5403–5413. [\[CrossRef\]](#)
111. Main, R.; Cho, M.A.; Mathieu, R.; O’Kennedy, M.M.; Ramoelo, A.; Koch, S. An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 751–761. [\[CrossRef\]](#)
112. Guyot, G.; Baret, F. Utilisation de la haute resolution spectrale pour suivre l’état des couverts vegetaux. In Proceedings of the Spectral Signatures of Objects in Remote Sensing, Aussois (Modane), France, 18–22 January 1988; p. 279.
113. Vogelmann, J.E.; Rock, B.N.; Moss, D.M. Red edge spectral measurements from sugar maple leaves. *TitleREMOTE Sens.* **1993**, *14*, 1563–1575. [\[CrossRef\]](#)
114. Reyniers, M.; Walvoort, D.J.J.; De Baardemaaker, J. A linear model to predict with a multi-spectral radiometer the amount of nitrogen in winter wheat. *Int. J. Remote Sens.* **2006**, *27*, 4159–4179. [\[CrossRef\]](#)
115. Moghimi, A.; Yang, C.; Marchetto, P.M. Ensemble feature selection for plant phenotyping: A journey from hyperspectral to multispectral imaging. *IEEE Access* **2018**, *6*, 56870–56884. [\[CrossRef\]](#)
116. Cai, J.; Luo, J.; Wang, S.; Yang, S. Feature selection in machine learning: A new perspective. *Neurocomputing* **2018**, *300*, 70–79. [\[CrossRef\]](#)

117. Johannes, M.; Brase, J.C.; Fröhlich, H.; Gade, S.; Gehrmann, M.; Fälth, M.; Sülthmann, H.; Beißbarth, T. Integration of pathway knowledge into a reweighted recursive feature elimination approach for risk stratification of cancer patients. *Bioinformatics* **2010**, *26*, 2136–2144. [CrossRef]
118. Zhang, C.; Li, Y.; Yu, Z.; Tian, F. Feature selection of power system transient stability assessment based on random forest and recursive feature elimination. In Proceedings of the 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xi'an, China, 25–28 October 2016; pp. 1264–1268.
119. Feng, L.; Li, Y.; Wang, Y.; Du, Q. Estimating hourly and continuous ground-level PM_{2.5} concentrations using an ensemble learning algorithm: The ST-stacking model. *Atmos. Environ.* **2020**, *223*, 117242. [CrossRef]
120. González Sánchez, A.; Frausto Solís, J.; Ojeda Bustamante, W. Predictive ability of machine learning methods for massive crop yield prediction. *SJAR* **2014**, *12*, 313–328. [CrossRef]
121. Jaikla, R.; Auephanwiriyakul, S.; Jintrawet, A. Rice yield prediction using a support vector regression method. In Proceedings of the 2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Krabi, Thailand, 14–17 May 2008; pp. 29–32.
122. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [CrossRef]
123. Awad, M.; Khanna, R. Support vector regression. In *Efficient Learning Machines*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 67–80.
124. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
125. Zhou, X.; Zhu, X.; Dong, Z.; Guo, W. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J.* **2016**, *4*, 212–219.
126. Wang, Y.; Wu, X.; Chen, Z.; Ren, F.; Feng, L.; Du, Q. Optimizing the Predictive Ability of Machine Learning Methods for Landslide Susceptibility Mapping Using SMOTE for Lishui City in Zhejiang Province, China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 368. [CrossRef]
127. Yang, Q.; Shi, L.; Han, J.; Zha, Y.; Zhu, P. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Res.* **2019**, *235*, 142–153. [CrossRef]
128. Dayananda, S.; Astor, T.; Wijesingha, J.; Chickadibburahalli Thimappa, S.; Dimba Chowdappa, H.; Nidamanuri, R.R.; Nautiyal, S.; Wachendorf, M. Multi-Temporal Monsoon Crop Biomass Estimation Using Hyperspectral Imaging. *Remote Sens.* **2019**, *11*, 1771. [CrossRef]
129. Wilkerson, S. Application of the Paired t-test. *XULaneXUS* **2008**, *5*, 7.
130. Hansen, P.M.; Schjoerring, J.K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* **2003**, *86*, 542–553. [CrossRef]
131. Cazenave, A.-B.; Shah, K.; Trammell, T.; Komp, M.; Hoffman, J.; Motes, C.M.; Monteros, M.J. High-Throughput Approaches for Phenotyping Alfalfa Germplasm under Abiotic Stress in the Field. *Plant Phenome J.* **2019**, *2*, 1–13. [CrossRef]
132. Cho, M.A.; Skidmore, A.; Corsi, F.; Van Wieren, S.E.; Sobhan, I. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 414–424. [CrossRef]
133. Clevers, J.G.P.W. Imaging spectrometry in agriculture-plant vitality and yield indicators. In *Imaging Spectrometry—A Tool for Environmental Observations*; Springer: Berlin/Heidelberg, Germany, 1994; pp. 193–219.
134. Gabriel, J.L.; Zarco-Tejada, P.J.; López-Herrera, P.J.; Pérez-Martín, E.; Alonso-Ayuso, M.; Quemada, M. Airborne and ground level sensors for monitoring nitrogen status in a maize crop. *Biosyst. Eng.* **2017**, *160*, 124–133. [CrossRef]
135. Frame, J.; Merrilees, D.W. The effect of tractor wheel passes on herbage production from diploid and tetraploid ryegrass swards. *Grass Forage Sci.* **1996**, *51*, 13–20. [CrossRef]
136. Meek, B.D.; Carter, L.M.; Garber, R.H.; Rechel, E.A.; DeTar, W.R.; Shatter, C.A. Regrowth and Yield of Alfalfa as Influenced by Wheel Traffic. Available online: http://works.bepress.com/william_detar/24/ (accessed on 3 March 2020).

