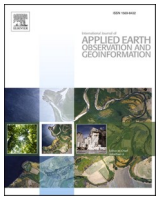
International Journal of Applied Earth Observation and Geoinformation 142 (2025) 104729

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Alfalfa stem count estimation using remote sensing imagery and machine learning on Google Earth Engine☆

Hazhir Bahrami [a,*], Karem Chokmani [a], Saeid Homayouni [a], Viacheslav I. Adamchuk [b], Md Saifuzzaman [b], Rami Albasha [c], Maxime Leduc [c]

[a] Centre Eau Terre Environnement, Institut National de la Recherche Scientifique, Québec, Canada
[b] Bioresource Engineering Department, McGill University, Macdonald Campus, Ste-Anne-de-Bellevue, QC, Canada
[c] My Forage Systems, Canada

ARTICLE INFO

Keywords:
Remote sensing
Alfalfa stem count
Machine learning
Winter mortality detection
Satellite data

ABSTRACT

Alfalfa (*Medicago sativa L.*), a perennial legume forage crop, is valued for its high yield and quality. However, its survival during winter can be affected by several factors, and its mortality significantly impacts alfalfa production, necessitating timely and spatially detailed monitoring. This study aims to propose a framework for estimating alfalfa stem density using satellite imagery and machine learning (ML) algorithms, which can lead to winter mortality detection early in the spring and provide a better understanding of potential total dry matter. Three ML models—support vector machine (SVM), random forest (RF), and extreme gradient boosting (XGB)—were applied to Harmonized Landsat Sentinel (Landsat only, which is HLSL30) and Sentinel-2 datasets, accessed via the Google Earth Engine (GEE) Python API. Two scenarios were evaluated: 1) single-date data, capturing satellite images within a 3-day time window to the date of field sample measurement, and 2) time-series data, in which three satellite images were collected for the measurements during the first growing cycle. Both classification and regression models were used in both scenarios to estimate and classify alfalfa stem density. ML Classification models categorized stem density into four groups (bare, low-density, medium-density, and high-density), achieving an accuracy of up to 85 % using Sentinel-2 data and 84 % using HLSL30 data. The results also indicated that alfalfa stem density can be estimated with an error of ~ ±6-9 stems/foot^2 (1 foot = 30.48 cm) using ML regression models. RF outperformed XGB and SVM in classification and regression tasks, showing superior accuracy in classifying density and lower root mean square error (RMSE) in estimating stem density. Our proposed framework model can offer valuable information to growers and decision-makers, enabling them to make timely and informed decisions.

1. Introduction

The global food supply is at risk due to population growth and climate change (Zhao et al., 2017). Global yield production of major cereal crops is expected to drop by 7 to 23 % by the end of the century if farming practices are not adapted to the changing climate (Rezaei et al., 2023). Climate change also impacts livestock productivity, in both direct and indirect ways, by affecting the quantity and quality of forage resources (Giridhar and Samireddypalle, 2015). By mid-century, climate change may enhance the forage crop productivity in the cold and humid regions of North America (Thivierge et al., 2023). This improvement is due to rising temperatures, higher atmospheric CO_2 levels, and an extended growing season compared to the 1990 – 2000 timeframe (Thivierge et al., 2023). However, increased winter thaws, less protective snow cover, and rising summer drought events may adversely affect winter survival and summer regrowth. Consequently, this may lead to a potential decline in the resilience of perennial forages gradually (Thivierge et al., 2023).

Alfalfa is one of the most significant crops among forage legumes (Kayad et al., 2016). It is referred to as the "queen of forage" owing to its high yield, nutritional attributes, and ability to sequester carbon (Saifuzzaman et al., 2022). It is grown in over 80 countries, covering a total area of 35 million hectares (Kayad et al., 2016). Alfalfa, or mixtures of alfalfa and grass, constitutes 44 % of the perennial forage areas in

☆ This article is part of a special issue entitled: 'Smart Ag with GeoInfo' published in International Journal of Applied Earth Observation and Geoinformation.

* Corresponding author.
E-mail address: hazhir.bahrami@inrs.ca (H. Bahrami).

https://doi.org/10.1016/j.jag.2025.104729
Received 5 February 2025; Received in revised form 8 July 2025; Accepted 9 July 2025
Available online 14 July 2025

Eastern Canada (Canada, 2021; Thivierge et al., 2023) and 20 % in the northeastern United States (Cherney et al., 2020). Alfalfa is a perennial crop, which indicates that its growing period spans several years under favorable growing conditions. It is cut several times annually, and typically, each growing cycle spans 30 to 35 days between mid-April and September in the Canadian croplands.

Nevertheless, climatic conditions such as freezing temperatures without protective snow cover, cyclical warm and freezing air temperatures, fall soil moisture, and the presence of ice sheets have been linked to decreased winter survival rates (Castonguay et al., 2006; McKenzie et al., 1988). These conditions occur in cold-temperature regions, particularly in Canada, the northern United States, and other areas with harsh winters. The resulting yield reduction can be a considerable concern for feedstock supplies (Sapkota et al., 2023). Consequently, detecting the damaged areas can assist growers, especially in Canada, where the growing season is relatively short, in preparing for specific agricultural practices while also enabling investors, consulting agencies, and decision-makers to proactively prepare for both importing and exporting of hay based on the predicted yield (Feng et al., 2020).

One standard method for evaluating alfalfa stand productivity potential and deciding whether to keep a damaged stand is to count the number of stems. Conventional approaches in assessing stem count, such as quadratic frame sampling, are time-consuming, challenging, and require much physical effort when collecting data in the field, and this may hinder the rapid sharing of knowledge in the field (Noland et al., 2018). Using these methods can pose significant challenges and may be impossible in large areas (Kayad et al., 2016). However, remote sensing (RS) technology enables the simultaneous and broad observation of plant growth with high timeliness and indirect touch.

RS allows non-destructive, efficient, and dynamic plant growth monitoring (Cai et al., 2018; Xie et al., 2008). Furthermore, RS technology has significantly contributed to the management of agricultural production, particularly in terms of monitoring crop distribution, estimating crop yield, assessing the impact of disasters, and other related factors (Atzberger, 2013; Meroni et al., 2013; Wang et al., 2023). RS systems typically generate a substantial amount of data due to the high spatial, spectral, radiometric, and temporal resolutions required for precision agriculture applications (Kamilaris et al., 2017). In the era of smart agriculture, traditional processing methods are unable to manage large data and meet the growing demands effectively (Pokhariyal et al., 2023). In this regard, machine learning (ML) has become a powerful method for effectively obtaining information from various sources.

ML, a subset of artificial intelligence, is a data analysis methodology that enables computer systems to automatically identify patterns from data without specific programming. Many thorough reviews emphasize the use of ML algorithms in agricultural management (Bahrami et al., 2022b; Chlingaryan et al., 2018; Liakos et al., 2018). ML is a practical approach that can offer a more accurate prediction of biochemical and biophysical characteristics of different crops (Bahrami et al., 2022a; Van Klompenburg et al., 2020). ML algorithms have been applied to predict the yield of several crops, including maize (Schwalbert et al., 2018), wheat (Zhang et al., 2020), and alfalfa (Feng et al., 2020), and other crops.

Previous studies have only focused on proposing alfalfa yield and some crop parameter models, most of them in smaller-scale or localized settings (Chen et al., 2024; Kayad et al., 2016). For example, Bahrami et al., (2025a) utilized Sentinel-2 satellite imagery to estimate alfalfa crop height. A few recent studies have utilized satellite or unmanned aerial vehicle (UAV)-acquired remote sensing images to build models for predicting alfalfa yield (Chandel et al., 2021; Dvorak et al., 2021; Feng et al., 2020). Echeverría et al. (2021) utilized Sentinel-2 data to estimate alfalfa fractional cover over 172 field sample points. Nevertheless, there is a lack of studies to accurately assess the number of alfalfa stems at the field, regional, or countrywide levels. Quantifying stem density per unit area is a straightforward technique for evaluating the productivity potential of an alfalfa crop. Undersander et al. (1998) demonstrated that the correlation between stem density and yield potential remains constant regardless of the age of the stand. To the best of our knowledge, there is only one research (Bahrami et al., 2025b) that has focused on alfalfa stem count using remote sensing imagery. However, Bahrami et al. (2025b) utilized proximal images, which cannot be applied on a large scale.

This research presents a comprehensive and practical framework for assessing alfalfa stem density and detecting winter mortality through field surveys and satellite data from Sentinel-2 and Harmonized Landsat Sentinel (Landsat part only, HLSL30). This study evaluated the potential of remote sensing data in estimating alfalfa stem density in four provinces in eastern Canada by applying various ML algorithms. The specific objectives of this study include: 1) to assess and compare the performance of vegetation indices and accuracy of various ML models to predict the alfalfa stem density and classify them over alfalfa fields during different growth stages, study areas, soil types, and topography conditions, 2) to use single-date and time-series remote sensing data and assess the advantages and disadvantages of each, 3) to generate alfalfa stem maps and assess the spatial variability in field productivity using potential stem maps, and 4) to provide farmers and decision-makers across the country with a precise and straight-forward framework to be capable of monitoring their alfalfa field as simple as possible.

2. Materials and methods

2.1. Datasets

2.1.1. Study area and field measurements

Field measurements were conducted on 597 alfalfa fields spread over four Canadian provinces—Nova Scotia, Quebec, Ontario, and Manitoba—using data collected over three years from 2021, 2022, and 2023 (see Table 1 for field numbers per province). Ground measurements included soil samples, stem counts, and crop heights, along with the image dataset (Fig. 1).

33 consultants and 192 producers participated in the field campaigns. In each field, a randomized design was used. Stem counts were taken in the spring and fall at each location, consisting of three data points referred to as landmarks (Fig. 2). A rectangular quadrat was used to measure the stem counts for each landmark position (Fig. 2). Field advisors measured the crop's height and the number of stems at three points in a triangular shape with a side of around two meters at each spot. Only alfalfa stems higher than 5 cm were counted. Next, we calculated the number of stems for the center of the triangle by taking the average of these three ground measurements (Fig. 2). In total, 22,664 measurements were collected during the three years. After calculating the mean values for the center points, 7445 points remained

2.1.2. Remotely sensed data

The satellite dataset used in this research (Sentinel-2 and HLSL30) was obtained using the Google Earth Engine (GEE) Python API. GEE (https://earthengine.google.com/), launched in 2010 by Google, is a cloud computing tool that allows geospatial analysis worldwide utilizing Google's infrastructure (Gorelick et al., 2017). The system consists of a multi-petabyte data catalog prepared for analysis, along with a high-performance, inherently parallel computing service (Gorelick et al., 2017). Both Sentinel-2 and HLSL30 data are freely available in GEE.

Table 1
The number of fields for each year for each province in Canada.

Province	Year		
	2021	2022	2023
Manitoba	46	22	21
Nova Scotia	4	4	4
Ontario	15	9	14
Quebec	532	492	464

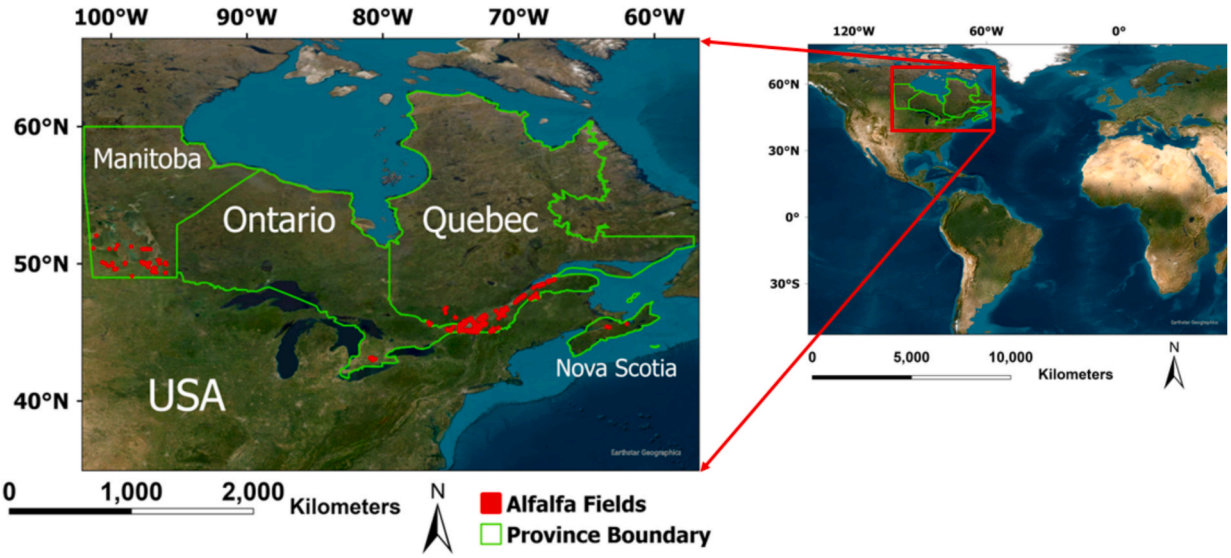


Fig. 1. Study area and data collection fields (red rectangles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

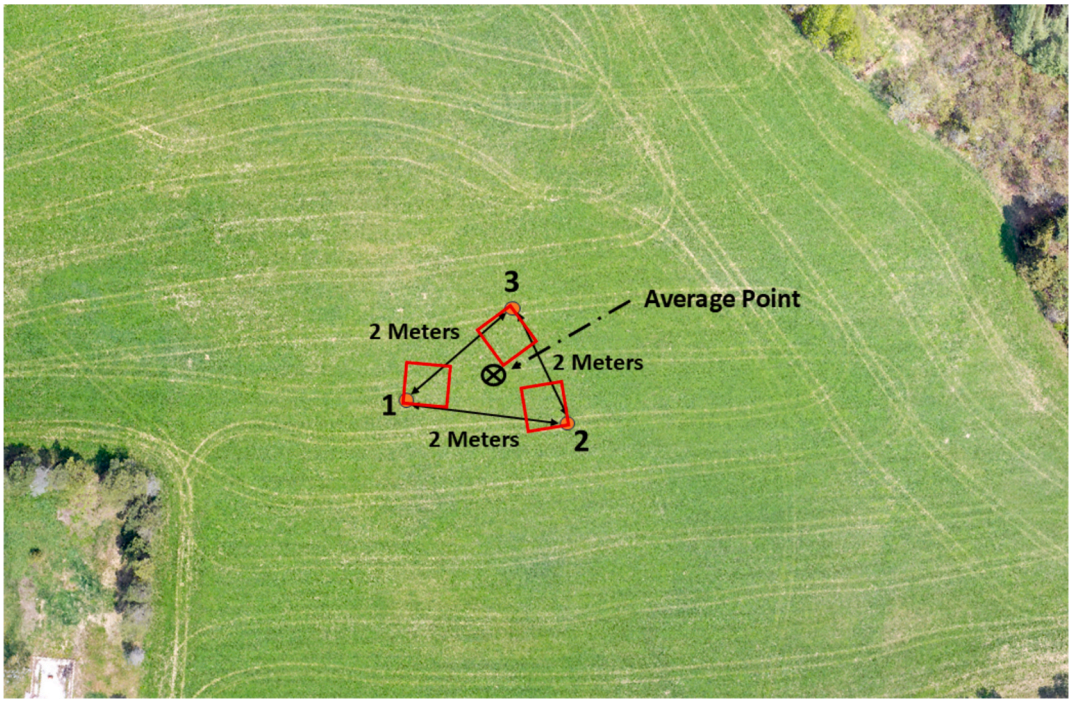


Fig. 2. The details of the protocol used in data collection, 3 sampling spots were placed in the field; measurements were taken within quadrats and placed at the corner of each landmark. Please note that the scale has been changed for better visualization.

The Sentinel-2 mission comprises two satellites: Sentinel-2A and Sentinel-2B. Sentinel-2 is a multispectral sensor that has thirteen bands, eight bands in the VNIR spectrum, and five bands in the SWIR range. With one satellite, the revisit time is 10 days, and it improves to 5 days when utilizing both satellites. The images utilized in this investigation were Sentinel-2 Level-2A data, which have undergone atmospheric correction and are accessible in GEE. All 10-m and 20-m spatial resolution bands were used in this study. Following the Sentinel-2 band reflectance extraction, several VIs have been extracted (Table 2).

The Harmonized Landsat and Sentinel-2 (HLS) project is a NASA endeavor designed to provide a continuous surface reflectance record from the Operational Land Imager (OLI) and Multi-Spectral Instrument (MSI) on the Landsat-8/9 and Sentinel-2A/B satellites, respectively. The HLS project (Claverie et al., 2017) aims to standardize data from both satellite projects to facilitate their integrated utilization. The program seeks to eliminate biases between the two sensors regarding varying spectral band ranges and viewing geometries, ultimately aiming to achieve worldwide land surface coverage at a spatial resolution of 30 m with a temporal interval of 2–3 days (Falanga Bolognesi et al., 2020). We selected only the Landsat part to compare the results with those from the sole Sentinel-2 and determine which sensor better estimates alfalfa density. Bands blue, green, red, NIR, SWIR 1 and 2, and TIRS 1 and 2 were extracted from the HLSL30 data.

Any images of Sentinel-2 and HLSL30 data were extracted within three days of the date of ground measurements. No more preprocessing was applied to the Sentinel-2 and HLSL30 data downloaded from GEE. A

Table 2
The details of the vegetation indices (VIs) used in this study. A symbol (✓) was added to the HLSL30 columns, indicating that the VI could be calculated if the same bands are available in the HLSL30 data.

Vegetation Index	Formula	HLSL30	Ref	Abbreviation
Normalized Green Red Vegetation Index	$\frac{Green - Red}{Green + Red}$	✓	(Gitelson et al., 2002)	NGRVI
Visible Atmospheric Resistance Index	$\frac{Green - Red}{Green + Red - Blue}$	✓	(Gitelson et al., 2003)	VARI
Visible-band Difference Vegetation Index	$\frac{2Green - Blue - Red}{2Green + Blue + Red}$	✓	(Xiaoqin et al., 2015)	VDVI
Green–Red Ratio Index	$\frac{Green}{Red}$	✓	(Gamon and Surfus, 1999)	GRRI
Normalized Difference Vegetation Index	$\frac{NIR - Red}{NIR + Red}$	✓	(Rouse et al., 1974)	NDVI
Normalized Difference Index 45	$\frac{Red\_edge1 - Red}{Red\_edge1 + Red}$		(Delegido et al., 2011)	NDI45
Normalized Difference Water Index	$\frac{Green - NIR}{Green + NIR}$	✓	(McFeeters, 1996)	NDWI
Normalized Difference Red Edge	$\frac{Red\_edge4 - Red\_edge1}{Red\_edge4 + Red\_edge1}$		(Gitelson and Merzlyak, 1994)	NDRE
Soil Adjusted Vegetation Index	$\frac{1.5(NIR - Red)}{NIR + Red + 0.5}$	✓	(Huete, 1988)	SAVI
Modified Soil Adjusted Vegetation Index	$\frac{2NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - Red)}}{2}$	✓	(Qi et al., 1994)	MSAVI
Enhanced Vegetation Index	$\frac{2.5(NIR - Red)}{NIR + 6Red - 7.5Blue + 1}$	✓	(Huete et al., 1997)	EVI
Chlorophyll Vegetation Index	$\frac{NIR*Red}{Green^2}$		(Vincini and Frazzi, 2011)	CVI
Simple Ratio	$\frac{NIR}{Red}$	✓	(Jordan, 1969)	SR
Optimized Soil Adjusted Vegetation Index	$\frac{NIR - Red}{NIR + Red + 0.16}$	✓	(Qi et al., 1994)	OSAVI
Modified Chlorophyll Absorption in Reflectance Index	$(Red\_edge1 - Red) - 0.2*(Red\_edge1 - Green)*(Red\_edge1 - Red)$	✓	(Daughtry et al., 2000)	MCARI
Inverted Red-Edge Chlorophyll Index	$\frac{Red\_edge3 - Red}{Red\_edge1/Red\_edge2}$		(Frampton et al., 2013)	IRECI

Blue: Band 2 in Sentinel-2 and HLSL30; Green: Band 3 in Sentinel-2 and HLSL30; Red: Band 4 in Sentinel-2 and HLSL30; NIR: Band 8 in Sentinel-2 and Band 5 in HLSL30; Red_edge1: Band5 in Sentinel-2; Red_edge2: Band6 in Sentinel-2; Red_edge3: Band7 in Sentinel-2; Red_edge4: Band8A in Sentinel-2.

threshold of 15 % was used for the maximum allowed cloud coverage percentage of the images. Cloud and cloud shadow masking were applied to the images. For Sensintel-2, cloudy and/or cloud shadow pixels were masked out using the QA60 bitmask and scene classification layer (SCL) bands that contain cloud information. For the HLSL30 data, cloud and cloud shadow removal was conducted using the Fmask band. Since the cloud masking does not remove all cloud and/or cloud shadow pixels, the images were checked one by one to ensure the high quality of data.

The details of the VIs extracted from Sentinel-2 and HLSL30 utilized in this study are shown in Table 2.

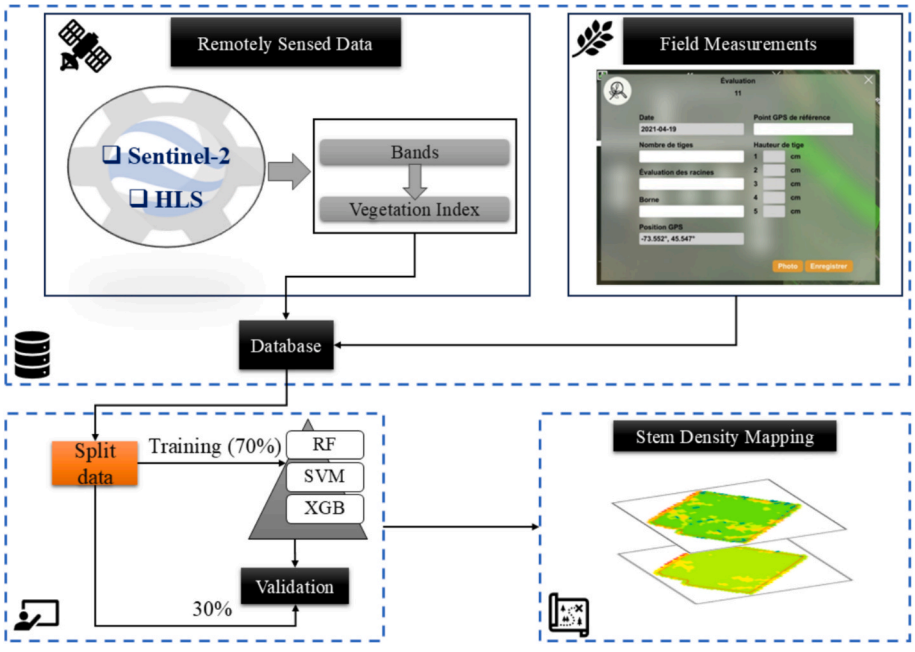


Fig. 3. The flowchart of the proposed methodology in this study.

2.2. General workflow

We utilized various ML regression and classification algorithms, 1) classification to classify stem densities into different classes, from bare to high-density, and 2) regression to estimate the stem density values. Although the stem density is a continuous value and regression is more suitable in this case, the rationale behind training classification models was to assess how the model performs across different classes. Two scenarios were considered to estimate/classify stem densities and detect winter mortality. In the first scenario, single-date remote sensing images within a 3-day time window of the ground measurement date were gathered for each of the alfalfa stem measurements during all growing cycles, and VIS were then calculated (Fig. 3). If there were no satellite measurements within a 3-day time window of the ground measurement date, or cloud and cloud shadows affected the area close to the sampling location, that measurement was eliminated.

After preprocessing and according to the timeline we considered, 2052 and 1747 samples were available for the Sentinel-2 and HLSL30 data, respectively, for the single-date scenario. To eliminate any potential inaccuracies in the data, we used the Python CleanLab package (Zhou et al., 2023). This package assists in cleaning data and labels by automatically identifying problems in an ML dataset. The Sentinel-2 and HLSL30 datasets had 1692 and 1528 points remaining, respectively, after applying CleanLab.

Based on agronomic guidelines (Undersander et al., 1998), stem densities are divided into three categories: 1) Alfalfa yield will diminish if the plant density falls below 40 stems/foot^2, 2) A stem count between 40 and 55 stems/foot^2 produces a maximum yield, and 3) A stem count higher than 55 has no effect on production (Undersander et al., 1998). We consequently split the stem counts into four classes for the classification models: 1) bare (count $<$ 5 stems/foot^2), 2) low-density (5 $<$ count $<$ 35 stems/foot^2), 3) medium-density (35 $<$ count $<$ 55 stems/foot^2), and 4) high-density (count $\geq$ 55 stems/foot^2).

The cleaned datasets were then split into training and validation sets. For regression models, stratification was performed by year to ensure temporal balance: 70 % of the data from each year were allocated to the training set, while the remaining 30 % were reserved for validation. For classification tasks, the data were stratified based on class labels to maintain class balance across splits. This resulted in 1184 training and 508 validation samples for Sentinel-2, and 1069 training and 459 validation samples for HLSL30.

Fig. 4 depicts the histogram representing the distribution of ground measurements for the stem count value using Sentinel-2 data in the first scenario.

In the second scenario, satellite measurements were extracted at the beginning, middle, and end of the first growing cycle. This is because the number of stems typically does not change throughout a single growing cycle. As previously mentioned, alfalfa is typically cut multiple times annually. Consequently, the term "first growing cycle" refers to the period from the beginning of the growing season in the spring to the first harvest. All in-situ measurements with at least three satellite measurements during the first growing cycle were considered. For each satellite measurement, all 10-m and 20-m spatial resolution bands, along with the VIs listed in Table 2, were extracted. The time-series data were subsequently fed into ML models to estimate and categorize the number of stems. Only Sentinel-2 multispectral data were used in this scenario, as the temporal resolution of HLSL30 in GEE data did not adequately cover the timeframe considered in this scenario. In total, 1803 in-situ measurements were available. After checking the satellite images and their corresponding dates one by one, 409 points were eliminated due to the potential presence of clouds and/or cloud shadows, or the date of availability of the satellite images was biased. In total, 1389 in-situ measurements were available; however, after applying the CleanLab, 847 points remained. Like the first scenario, the second scenario used the same stratifying technique to split the data: 70 % of the data was chosen for training, while 30 % was randomly selected for validation. For training and validation, 592 and 255 samples were available in the second scenario, respectively.

The importance of each feature has also been analyzed in this scenario. This allowed for defining the most informative features to assess alfalfa stem count. The parameters of each ML method were tuned using GridSearch cross-validation (GridSearchCV). Scikit-learn (Pedregosa et al., 2011) has an existing function called GridSearchCV. Finding the ideal values for the parameters in each model involves fine-tuning the model's hyperparameters using GridSearchCV. A cross-validation value of 5 has been used for this study. This means that 5-fold cross-validation is what we are utilizing for training. Five equal—or nearly equal—parts are randomly selected from the dataset. Four parts are used to train the model, while the remaining part is used for testing. Five times through this process, a different part is employed as the test set each time.

2.3. Machine learning methods

2.3.1. Random forests

Random Forests (RF) are ensemble learning models developed by Breiman (2001) utilized for classification and regression tasks. Ensemble methods employ several learning mechanisms to enhance performance. Boosting and bagging are the major ensemble learning techniques. Boosting involves creating a sequence of models, wherein each model aims to fix the errors of the previous one. Multiple base models are

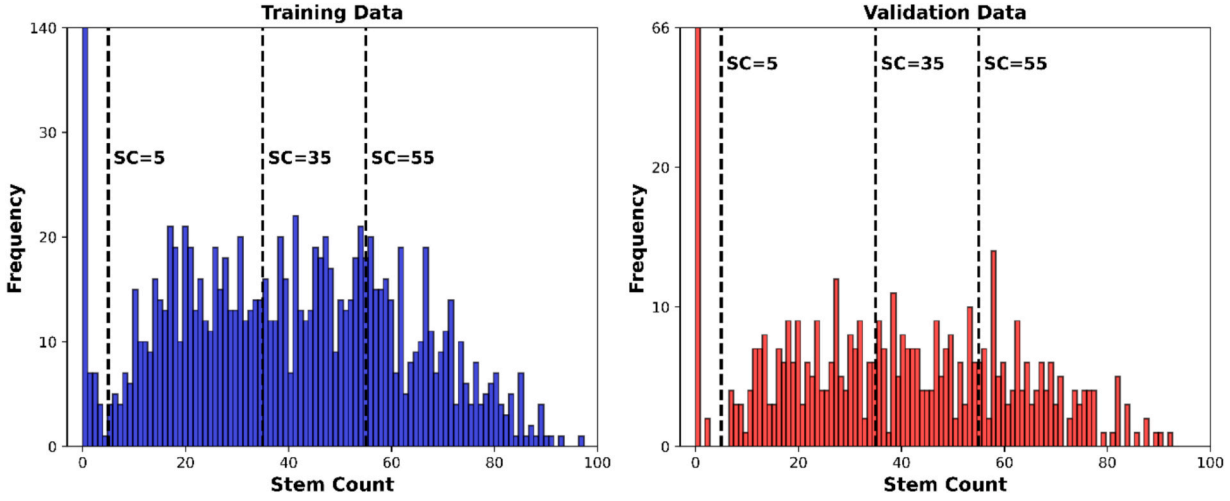


Fig. 4. The histogram of the stem count data in the training data (left) and validation data (right). (SC = Stem Count).

independently fitted during the bagging process, resulting in a more stable composite model with reduced variance. In RF, which is a successful bagging ensemble approach, a collection of estimators is integrated to provide a more precise model. The RF algorithm mitigates model variance by averaging the outputs of all decision trees (Dangeti, 2017). The details of the grid parameters used in this study for RF tuning are shown in Table 3.

2.3.2. Support vector machine

The Support Vector Machine (SVM) model was introduced by Cortes and Vapnik (1995). It is a robust and extensively utilized kernel-based ML algorithm for classification tasks. The SVM model can be adapted for regression problems (Awad et al., 2015). In SVR, the objective is to find a function f(x) that deviates from the targets by no more than epsilon. A flexible tube is constructed around the estimation function by employing SVR, thereby disregarding absolute error amounts that fall below a predetermined threshold. Points within the tube, whether above or below the prediction function, suffer no penalties, whereas points outside the tube are subject to penalties. The details of the grid parameters used in this study for SVM tuning can be seen in Table 4.

2.3.3. Extreme gradient boosting

XGB (Brownlee, 2016) is a widely utilized implementation of gradient boosting, initially created by Tianqi Chen in 2001 as a research work. The approach utilizes a gradient-boosting architecture and functions as an ensemble machine-learning algorithm. XGB improved a machine learning model's performance, speed, adaptability, and efficiency. Table 5 shows the GridSearch parameters utilized for tuning the XGB hyperparameters.

2.4. Accuracy assessment

Confusion matrices are insightful and detailed techniques to assess the performance of classification algorithms (Sathyanarayanan and Tantri, 2024). A confusion matrix is a two-by-two table that indicates the counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

The recall metric evaluates how well the model can identify positive samples, making detecting and correcting any overlooked positive samples crucial. Recall is computed according to (Equation (1). A high recall indicates that the model has a strong ability to recognize positive instances accurately. The precision of a model is defined by its positive detection accuracy and is calculated using Equation (2). High precision means that the model has a limited frequency of false positives. The F1-score (Equation (3) is a harmonic mean of recall and precision, which is valuable for considering both false positives and false negatives. Finally, accuracy is a metric that quantifies the proportion of accurately identified instances within the entire dataset, computed as per Equation (4).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{FP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{F1} - \text{score} = 2^{*}\frac{\text{Precision}^{*}\text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2^{*}\text{TP}}{2^{*}\text{TP} + \text{FP} + \text{FN}} \tag{3}$$

Table 3
The detail of the RF parameters optimized in this study using GridSearchCV.

Parameters	Description	Grid Search Values
n_estimators	Number of trees in the forest	10, 30, 50, 100
max_depth	Maximum depth of the trees	3, 4, 5, 6
max_features	The number of features to consider when looking for the best split	3, 5, 10

Table 4
The details of the SVM parameters optimized in this study using GridSearchCV.

Parameters	Description	Grid Search Values
Kernel	type of kernel that the algorithm employs.	'rbf', 'poly,' 'linear'
Gamma	Kernel coefficient for 'rbf', 'poly', and 'sigmoid.'	0.0001, 0.001, 0.05, 0.01, 0.05, 0.1, 0.5
C	Penalty parameter	1, 10, 50, 100, 200
Degree	The polynomial kernel function's degree	2, 3, 4
epsilon	Specifies the epsilon-tube within which no penalty for the points predicted within a distance epsilon from the actual value	0.05, 0.1, 0.2, 0.3, 0.5

Table 5
The details of XGB parameters optimized in this study using GridSearchCV.

Parameters	Description	Grid Search Values
learning_rate	Shrinks the contribution of each tree	0.001, 0.05, 0.01, 0.1, 0.2, 0.3, 0.5
n_estimators	The number of boosting stages to conduct.	10, 30, 50, 100
max_depth	Limits the number of nodes in the tree.	3, 4, 5, 7, 10

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{4}$$

For regression models, several criteria, including root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2), were used in this study to evaluate the performance of each machine learning model, calculated respectively in Equations (5), 6, and 7):

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\widehat{y}_i - y_i)^2}{n}} \tag{5}$$

$$\text{MAE} = \frac{\sum_{i=1}^{n}|\widehat{y}_i - y_i|}{n} \tag{6}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\widehat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\widehat{y}_i - \overline{y})^2} \tag{7}$$

where $\widehat{y}_i$ and y_i are respectively the observed and estimated i^{th} value, and n is the number of observations.

RMSE offers a measurable evaluation of the distance value, i.e., the difference between the observed and estimated values. The use of squared values makes RMSE sensitive to outliers. We also used MAE, which is less sensitive to extreme values, making it useful when pursuing an indicator that is less vulnerable to outliers. Finally, R^2 measures trend agreement, i.e., evaluating whether the estimated variable values increase with those observed and vice versa.

3. Results

3.1. Single-date stem count

3.1.1. Feature importance

Fig. 5 illustrates the feature importance results for RF and XGB using Sentinel-2 data. NDVI and OSAVI hold the highest RF and XGB feature importance positions, respectively. The results indicate that the normalized difference of the red, NIR, and red edge bands yields greater importance. VDVI ranks third among the important features for both RF and XGB, indicating that an index utilizing solely RGB bands can be crucial.

XGB feature importance results, like RF feature importance, indicate that NDVI, OSAVI, and VDVI are the most important among all the VIs and bands using Sentinel-2 data. In XGB feature importance, NDI45 and

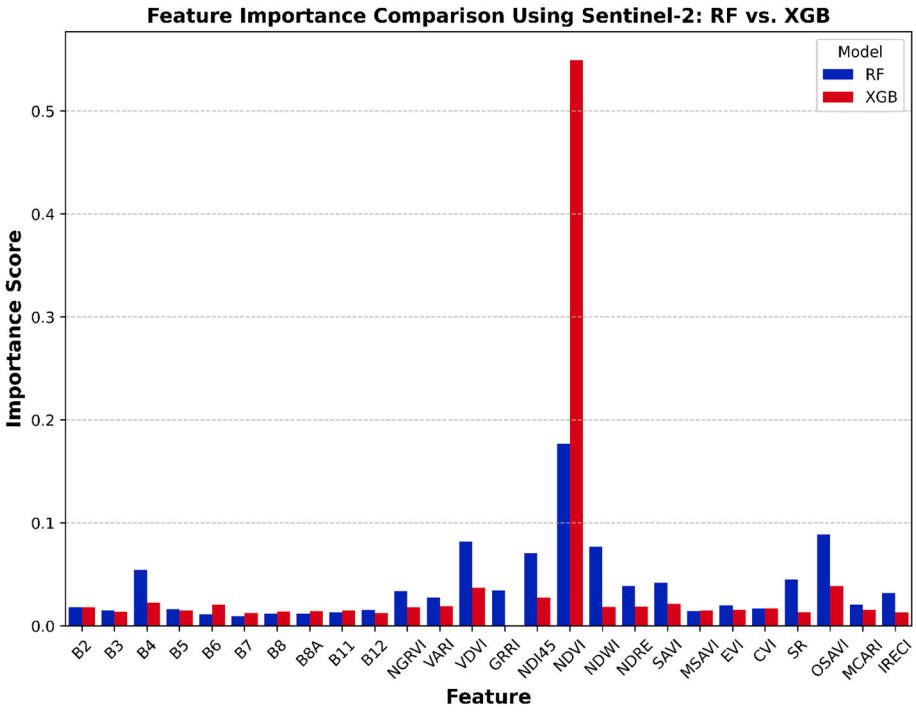


Fig. 5. The chart of feature importance using RF (blue) and XGB (red) using Sentinel-2 data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

B4 are ranked third and fourth, indicating that RGB VIs significantly contribute to the XGB model.

For HLSL30 data, NDVI and OSAVI had the highest ranks of feature importance for both RF and XGB, similar to the results obtained with Sentinel-2 data (Fig. 6). SAVI, NDWI, and EVI ranked next in RF, while NGRVI, VARI, and OSAVI ranked next in XGB.

3.1.2. *Performance evaluation of alfalfa stem density classification*

The results of ML algorithms to estimate the stem classes using Sentinel-2 data are shown in Table 6. RF achieved 89 % accuracy on the training dataset, while XGB and SVM reached 87 % and 85 % accuracy, respectively. RF has the highest accuracy for validation data, achieving an overall accuracy of 85 %. XGB and SVM ranked second and third, with an overall accuracy of 84 % and 83 % for the validation data, respectively.

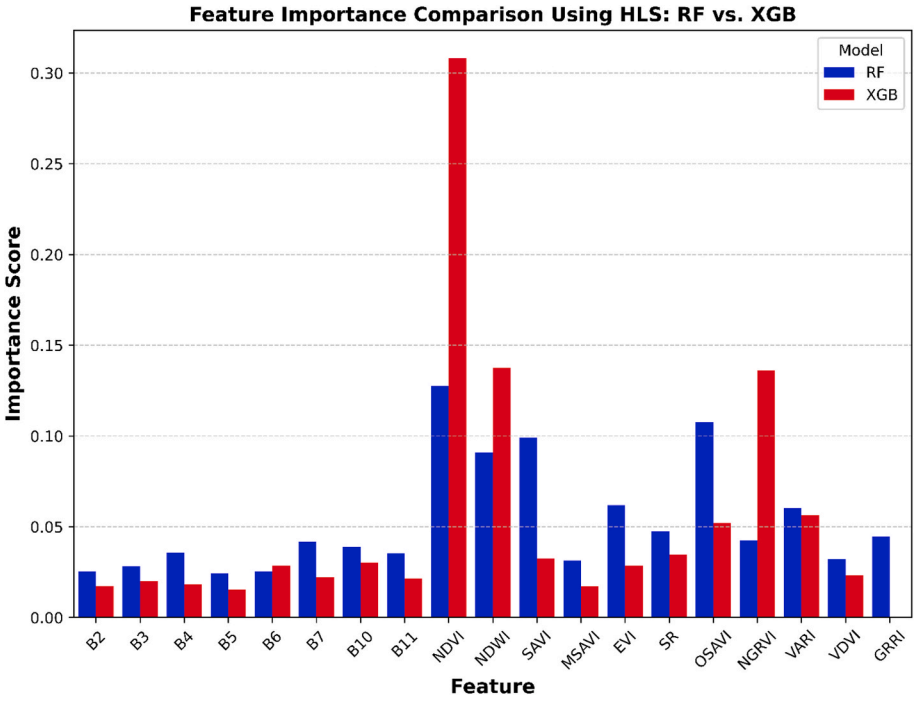


Fig. 6. The chart of feature importance using RF (blue) and XGB (red), using HLSL30 data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 6
Evaluation metrics of the ML algorithms in estimating the number of stems using single-date Sentinel-2 data.

		Sentinel-2					
		RF		XGB		SVM	
		Training	Validation	Training	Validation	Training	Validation
Bare	Precision	0.99	0.97	0.99	0.96	0.97	0.93
	Recall	0.96	1.00	0.96	1.00	0.96	1.00
	F1-score	0.98	0.99	0.97	0.98	0.97	0.96
Low-density	Precision	0.96	0.91	0.94	0.90	0.95	0.93
	Recall	0.91	0.87	0.90	0.85	0.87	0.81
	F1-score	0.93	0.89	0.92	0.88	0.91	0.86
Medium-density	Precision	0.75	0.69	0.73	0.68	0.72	0.67
	Recall	0.90	0.81	0.88	0.83	0.81	0.79
	F1-score	0.82	0.75	0.80	0.75	0.76	0.73
High-density	Precision	0.92	0.90	0.91	0.93	0.83	0.85
	Recall	0.80	0.77	0.77	0.74	0.83	0.81
	F1-score	0.86	0.83	0.83	0.83	0.83	0.83
Total	Accuracy	**0.89**	**0.85**	0.87	0.84	0.85	0.83

In the bare class, all models showed great potential in accurately classifying the classes compared to other classes in training and validation datasets. RF showed the highest accuracy in terms of F1–score, with values of 99 % for the validation dataset. For the low-density class, the F1-score for the training data was 93 %, but it was 89 % for the validation data. In the high-density class, the F1-score for the training data in RF surpassed that of XGB and SVM, although the validation F1-score for all models was 83 %. Based on the comparison between training and validation evaluation metrics, there is no sign of overfitting in any of the ML models used in this study utilizing Sentinel-2 data.

For the HLSL30 dataset, RF and XGB performed the best in the validation data in terms of overall accuracy (Table 7). The accuracy for the validation data was 84 % in RF and XGB, whereas it was 79 % in SVM. In the training data, the XGB model outperformed the other two models, achieving an overall accuracy of 100 % over all classes. This demonstrates that the XGB model has been overfitted to the training data. XGB outperformed the other two models for all classes in terms of F1-score over the validation dataset. The SVM model had the lowest rank in terms of overall accuracy in both the training and validation datasets.

All ML models demonstrated remarkable accuracy in classifying stem classes. The RF demonstrated slightly higher accuracy for both the HLSL30 and Sentinel-2 datasets. XGB held the second position, while SVM occupied the final place.

Fig. 7 displays the normalized confusion matrix by row for the validation datasets for Sentinel-2 (left column) and HLSL30 (right column) datasets. Minimal confusion is observed between bare and low-density classes, indicating a limited number of samples originally categorized as bare that were misclassified as low-density or vice versa. The most significant confusion across nearly all models was observed between the medium-density and high-density classes. Analysis of HLSL30 data revealed that the high-density class was identified with a higher accuracy; however, the utilization of Sentinel-2 data indicated that the medium-density class was classified with higher accuracy. The most significant confusion arises among pixels labeled as medium-density, which the model erroneously categorized as high-density across all models.

3.1.3. Performance evaluation of alfalfa stem density estimation

Fig. 8 depicts the results of different ML regression algorithms on validation data using Sentinel-2 data on the left and HLSL30 data on the right. The accuracy of RF and XGB for the validation data using Sentinel-2 was approximately similar. The R^2 value for both RF and XGB was 0.89. RF and XGB slightly surpassed SVR using Sentinel-2 data. Using the Sentinel-2 dataset, the value of RMSE for RF (7.83 stems/foot^2) was slightly better than those for XGB (7.88 stems/foot^2) and SVR (8.66 stems/foot^2). However, the value of MAE with XGB (6.03 stems/ft^2) was lower than those obtained with RF (6.06 stems/foot^2) and SVR (6.93 stems/foot^2). The SVR, however, had a lower R^2 (R^2 of 0.83) compared to RF and XGB. RF slightly outperformed XGB in terms of RMSE and MAE. Saturation in SVR can be seen in values above ~ 75 stems in Sentinel-2, while the saturation value is higher in RF and XGB.

Utilizing the HLSL30 dataset, like the Sentinel-2 dataset, RF and XGB had the same R^2 value (R^2 of 0.89). SVR ranked third with an R^2 value of 0.83. The MAE for XGB was approximately six stems. RF and XGB demonstrated superior performance using HLSL30 while analyzing validation data. Saturation in HLSL30 data becomes increasingly apparent as the values approach approximately 80 stems. SVR showed

Table 7
The results of various ML algorithms in estimating the number of stems using single-date HLSL30 data.

		HLSL30					
		RF		XGB		SVM	
		Training	Validation	Training	Validation	Training	Validation
Bare	Precision	0.94	0.93	1.00	0.94	0.96	0.92
	Recall	0.93	0.92	1.00	0.95	0.94	0.90
	F1-score	0.93	0.93	1.00	0.94	0.95	0.91
Low-density	Precision	0.86	0.87	1.00	0.89	0.85	0.85
	Recall	0.92	0.90	1.00	0.88	0.91	0.90
	F1-score	0.89	0.88	1.00	0.89	0.88	0.88
Medium-density	Precision	0.80	0.75	1.00	0.74	0.76	0.70
	Recall	0.70	0.74	1.00	0.76	0.66	0.61
	F1-score	0.75	0.75	1.00	0.75	0.71	0.65
High-density	Precision	0.83	0.84	1.00	0.83	0.81	0.73
	Recall	0.87	0.81	1.00	0.82	0.86	0.78
	F1-score	0.85	0.82	1.00	0.83	0.83	0.76
Total	Accuracy	0.85	0.84	**1.00**	**0.84**	0.83	0.79

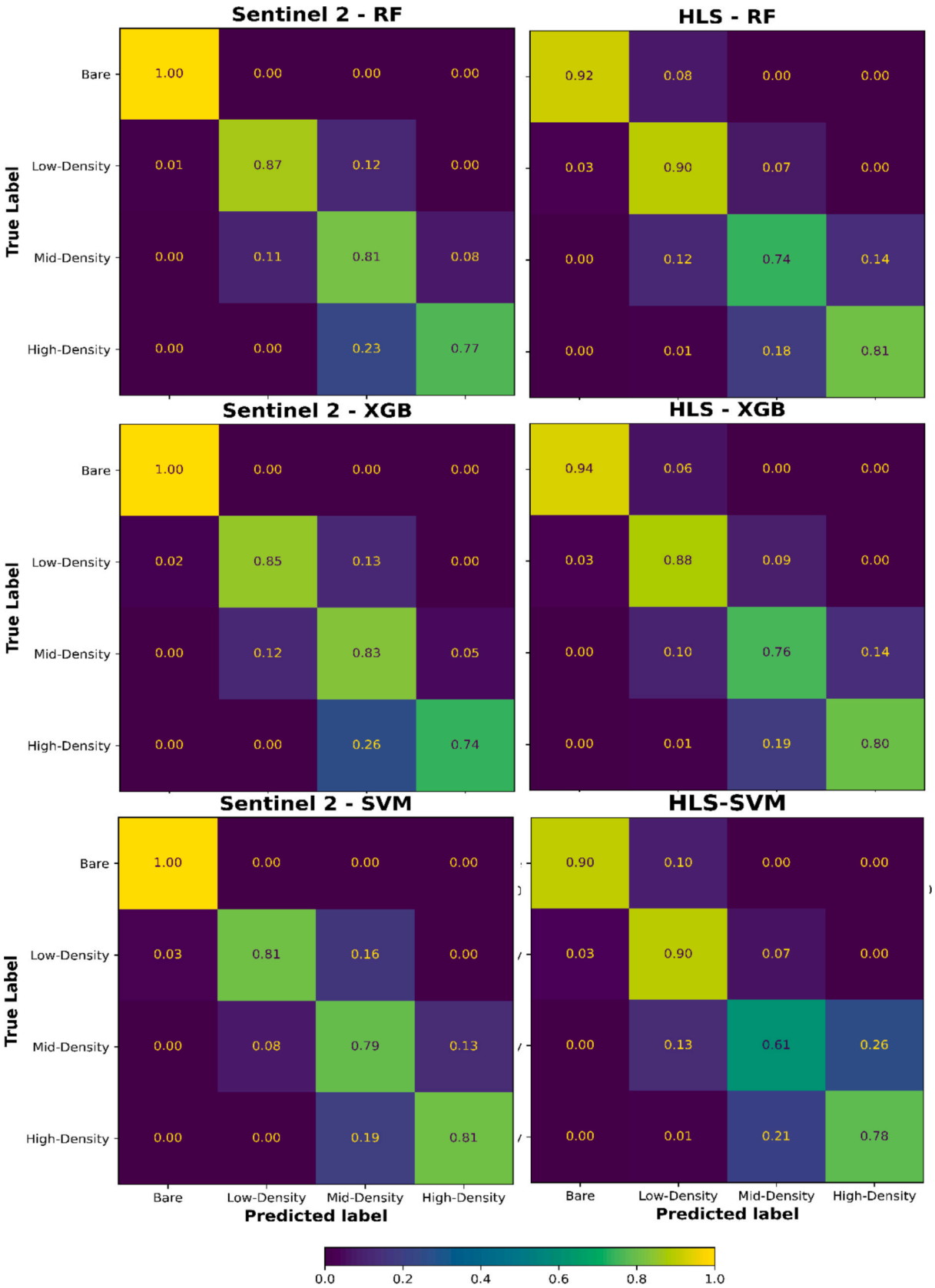


Fig. 7. The confusion matrices of validation data using Sentinel 2 data (left column) and HLSL30 data (right column) using RF, XGB, and SVM.

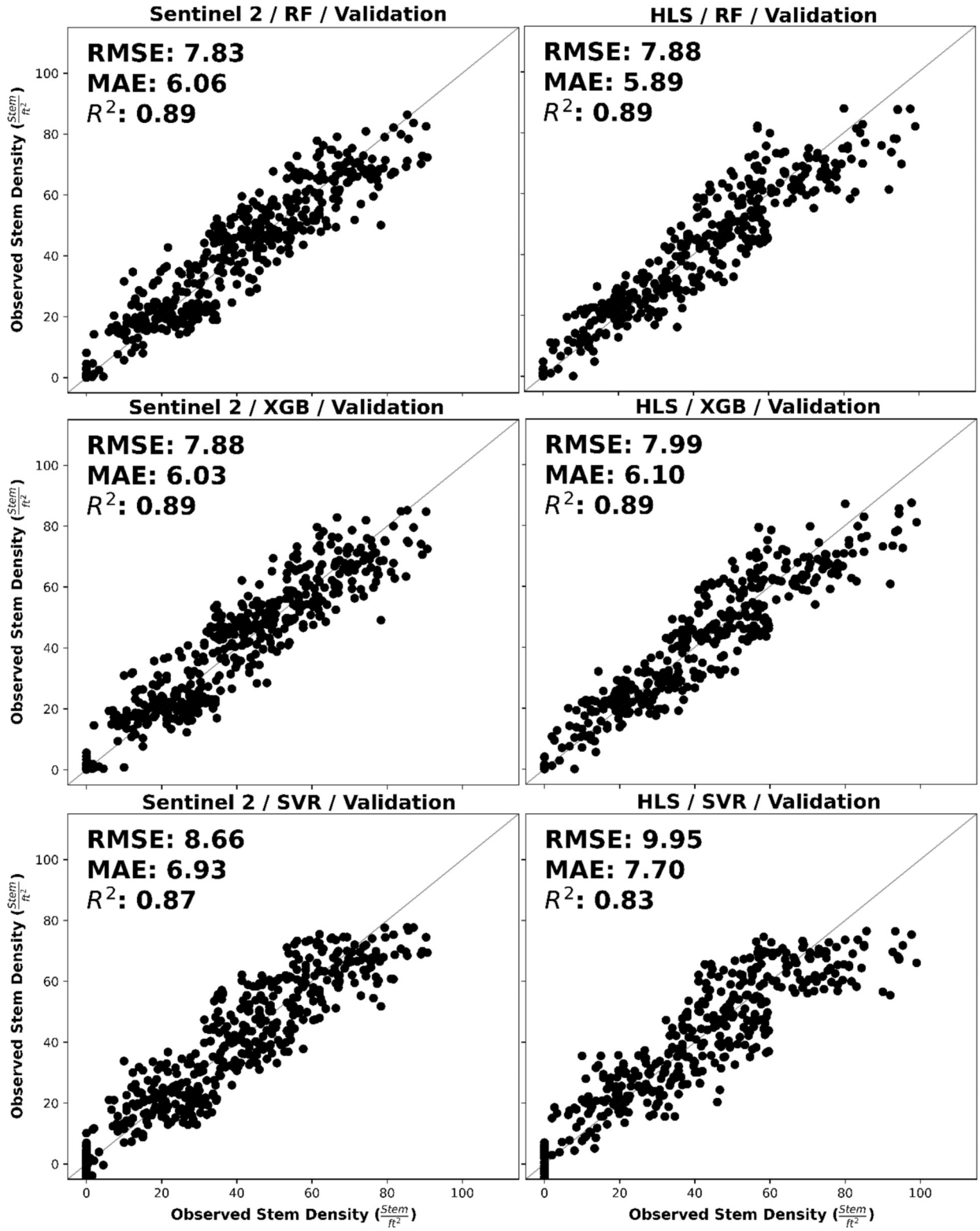


Fig. 8. The scatterplots between estimation and observation of various ML regression algorithms to estimate stem count on validation data using Sentinel-2 (left column) and HLSL30 (right column) data.

unusual behaviors in the sparse and low-density alfalfa region. The negative numbers estimated by SVR may result from normalization, as normalization was performed solely on the training data.

3.1.4. Stem count and stem class mapping

Fig. 9 shows the mapping of stems using both single-date regression and classification models. Using the RF regression and classification model, we acquired Sentinel-2 data from GEE due to its resolution and predicted the number of stems and stem classes on various dates. Three images that were taken before the harvest were utilized for mapping analysis. The findings indicated that the models may effectively project alfalfa distribution in future circumstances. As it is shown in Fig. 9-a and

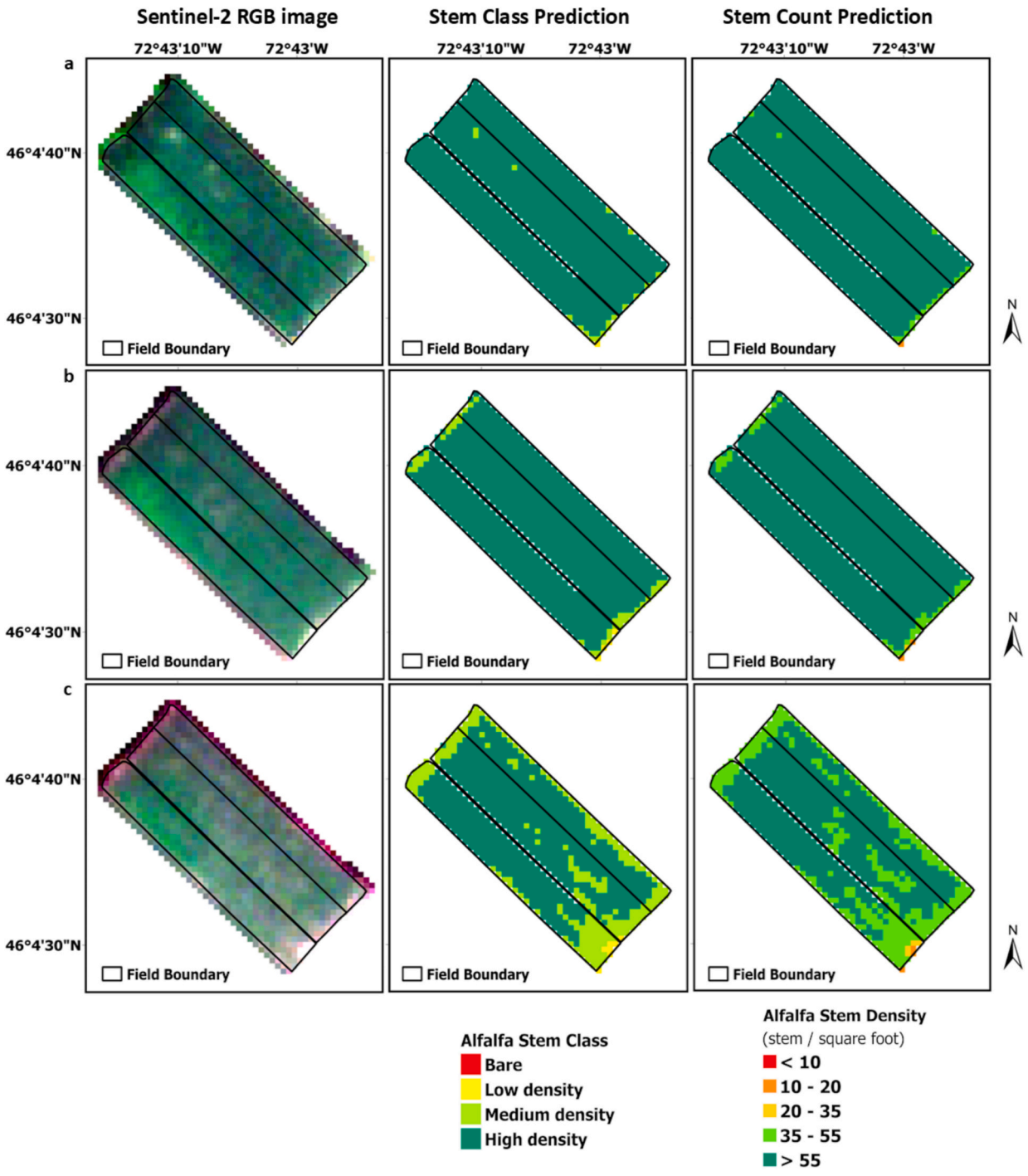


Fig. 9. Comparison of RGB images, stem class predictions, and stem count estimations for three dates: (a) May 27, 2023, (b) June 21, 2023, and (c) September 24, 2023. Each row corresponds to a specific date, with columns showing RGB images (left), stem class predictions (center), and stem count estimations (right).

-b, the quantity of stems remained relatively stable at the end of the first and the second growing cycles. Nonetheless, the mapping of stem quantities indicated a reduction of high-density class (more than 55 stems in the regression) during the third growing cycle (Fig. 9-c). Using this model, if a pixel is classified as bare in the classification model or the value of the regression model is close to zero, that location must be considered winter mortality, as the stem density does not satisfy yield potential.

3.2. Time-series stem count

3.2.1. Performance analysis of classifying and estimating alfalfa stem density

Like the results with single-date data, the RF outperformed other models in regression (Fig. 10). RF enabled the estimation of stems by utilizing three satellite images per point, achieving an RMSE of ~ 10 stems. The results indicated that employing RF and XGB for classification may predict the stem classes with approximately 82 % overall accuracy in the validation dataset. The SVM exhibited lower overall accuracy compared to RF and XGB, with an overall accuracy of 80 %.

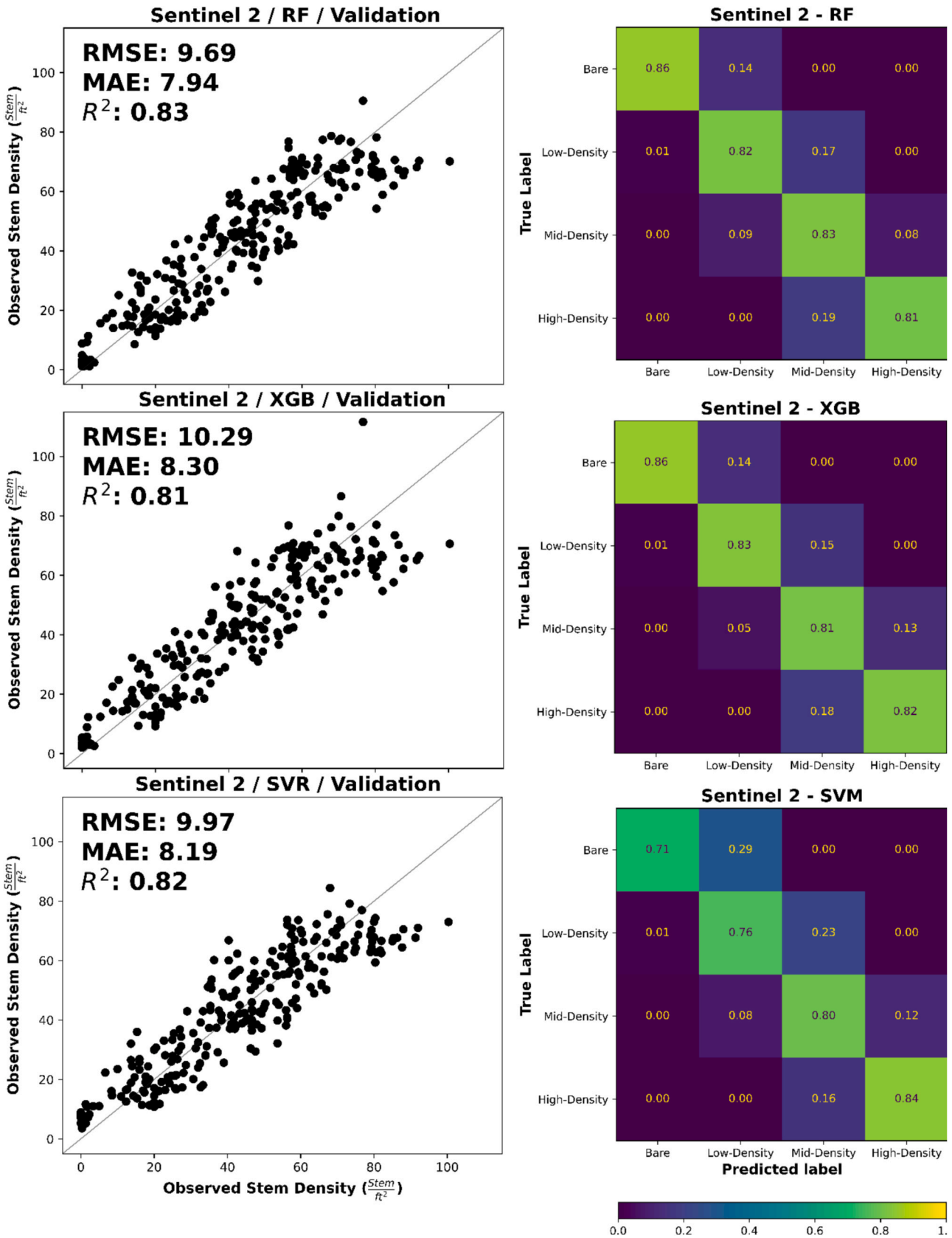


Fig. 10. The scatterplots between estimation and observation validation data using multi-temporal data (left column), and confusion matrix of validation data (right column) using Sentinel-2 data and XGB model.

Despite SVM exhibiting relatively high performance, having a few misclassified samples between medium- and high-density classes, its performance distinguishing the bare class from low-density and low-density from medium-density was lower. Despite the presence of misclassified pixels among classes in the classification models, these pixels probably fall around the threshold values. In regression models, nearly all exhibited comparable performance on the validation data, with an MAE of approximately eight stems. The R^2 score for XGB was 0.81, whereas it was 0.82 for SVR and 0.83 for RF. RF somewhat surpassed other models regarding RMSE and MAE, as the RMSE and MAE were 9.69 stems/foot^2 and 7.94 stems/foot^2 for RF, respectively, while they were 10.29 stems/foot^2 and 8.30 stems/foot^2 for XGB and 9.97 stems/foot^2 and 8.19 stems/foot^2 for SVR. All models show some saturation in the stem values of higher than ~ 80.

3.2.2. *Stem count and stem class mapping using Time-Series data*

Three different fields were selected for this section of the research. Three images for each field during the first growing cycle were collected

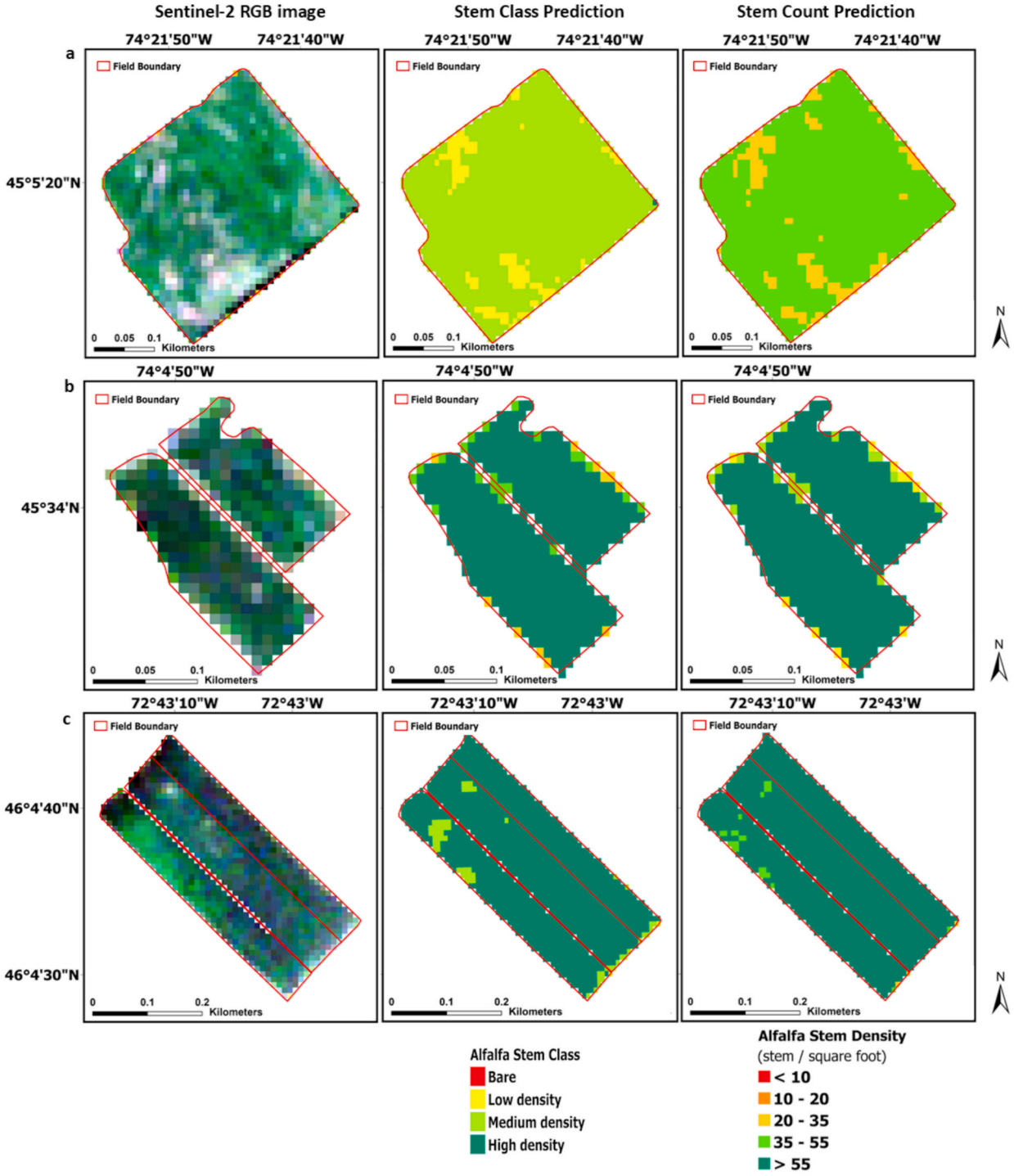


Fig. 11. Comparison of Sentinel-2 imagery, stem class classification, and stem count regression results for three different dates. Each row corresponds to a specific Sentinel-2 acquisition date (a: acquired on May 25, 2023; b: on May 30, 2023; and c: on May 27, 2023) with the Sentinel-2 RGB image (left), the stem class mapping from a classification model using multi-temporal data (center), and the estimated stem count per unit area from a regression model (right).

and fed into the ML algorithms. We quantified and classified the number of stems using the RF regression and classification models. As depicted in Fig. 10, the results indicated that the model's predictions closely matched the RGB images taken right before the first harvest. Also, the results indicated that the classification prediction of stem density closely matched the regression estimates. In this scenario, like the previous scenario, the maps can be utilized to detect winter mortality in the spring and to estimate the number of stems for the potential yield calculations. If a pixel is classified into the bare class in the classification model or close to zero in the regression, that location must be considered winter mortality as the stem density is low and does not satisfy yield potential. See (Fig. 11).

4. Discussion

Using satellite remote sensing data, this study proposed a robust and practical framework to count the number of alfalfa stems and detect winter mortality early in the season in four provinces in eastern Canada, including Quebec, Ontario, Manitoba, and Nova Scotia. We assessed the viability of Sentinel-2 and HLSL30 spectral bands and vegetation indices for estimating stem count variability in alfalfa fields using GEE. The performances of RF, SVM, and XGB ML algorithms for stem count and winter mortality detection were assessed and compared. The results showed that RF has great potential in estimating alfalfa stem density and, therefore, detecting winter mortality using remote sensing data.

Although there is a lack of research on estimating the alfalfa stem density using satellite data and ML algorithms, several studies have demonstrated the great potential of RF in predicting the crop biophysical parameters of other crops (Bahrami et al., 2022a; Bahrami et al., 2021; Kganyago et al., 2024; Kganyago et al., 2021). Bahrami et al. (2025b) estimated the number of alfalfa stems using proximal iPad images. They reported an RMSE of about 10 stems in their study. However, we achieved a higher MAE value of about 6–9 stems in our study. Additionally, it is impossible or difficult to apply proximal images on large-scale mapping. This study also evaluated the importance of each feature in the examined models. Our results indicated that NDVI was the most important feature in alfalfa stem density estimation. In a review study by Tedesco et al. (2022), NDVI was among the high-potential and highly used vegetation indices in alfalfa crop parameter estimation.

The findings indicated that both Sentinel-2 and HLS data have significant potential for estimating the intra-field variations of the number of stems. In various studies, the high potential of Sentinel-2 and HLS data in estimating biophysical parameters of multiple crops has been acknowledged (Dong et al., 2020; Kganyago et al., 2021; Qiao et al., 2024; Xu et al., 2025).

Given that a crucial purpose was to identify winter damage areas for farmers, the different scenarios were utilized to provide essential information. In one scenario, satellite images at various growing stages were employed to develop alfalfa stem count inversion models. This scenario can be applied as soon as the satellite image is available. Therefore, we can provide stem count maps once the satellite image is available, and thus avoid losing time. However, this approach yields varying stem counts for each satellite image, as it highly relies on the reflectance of various bands and vegetation indices. For example, two satellite images with only two- or three-day differences may result in different stem count maps since the reflectance in different bands in these two images differs. This causes problems as the number of stems remains constant during each growing period. Consequently, we cannot determine the optimal timing to utilize the model. The first scenario is practical for the early detection of dead areas in spring. This way, we may monitor the fields over time, and if a pixel consistently exhibits a low stem count, it indicates an issue at that particular area on the ground. Therefore, the model can warn the farmers and inform them of the exact location of potential dead or damaged areas in the early spring.

In an alternative scenario, we examined time-series satellite data. Three Sentinel-2 measurements have been considered for each sampling point in the first growing cycle. By extracting reflectance and vegetation indices and considering the consistency of stem count, we developed different ML models to predict the number of stems. RF showed significant capability in quantifying stem counts for time-series data. This model enables farmers to detect winter mortality early in spring, before the first harvest, and provides an accurate estimation of the number of stems, facilitating potential yield estimation. The sole limitation of this model is that it requires a minimum of three satellite images to generate its final assessment.

The results of this study indicate the potential for stem count estimates in alfalfa fields, which might help a deeper comprehension of yield distribution at the pixel level through remote sensing measurements. Based on a study by Dan Undersander et al. (2011), there is a linear relationship between the number of stems per square foot and the potential dry matter yield. Consequently, when the stem count for each pixel is provided, the potential total dry matter yield for each field can be calculated. We referred to potential dry matter, as the actual dry matter is dependent upon several factors, including weather, fertility, and drought, among others. Furthermore, stem count serves as a signal for farmers to determine whether to rotate their alfalfa fields.

The results of this study demonstrated slight saturation in the stem count models, especially for stem counts exceeding ~ 80. This may be attributed to either the saturation of VIs or the insufficient number of samples over 80 stems per foot. The problem of saturation of the vegetation index produces deviations in crop parameter estimation, such as LAI and yield estimation, and it has been widely stated in the literature (Gao et al., 2023; Wang et al., 2022). However, based on the literature, the alfalfa yield is not significantly affected by stem density exceeding ~ 55 stems/foot^2 (Maxime Leduc, 2019). Furthermore, as previously stated, a stem count exceeding 55 is regarded as highly dense, and this threshold does not necessitate crop rotation by the farmers. Consequently, this saturation will not constrain our models.

4.1. Limitations and future work

In this study, the Landsat part of the HLS data has only been used. Using both sensors could be of great importance in future work. The only setback of the HLS data is its coarse resolution. However, the temporal resolution can offer rich data since the temporal resolution could reach ~ 2 to 3 days, which would be a great asset, especially in Canada, where the growing cycles of alfalfa are relatively short. In that case, we would have more measurements, and the models would be more reliable. Having high temporal resolution data enables us to utilize other state-of-the-art models, such as the long short-term memory (LSTM) model, which can be more accurate than conventional ML models.

As mentioned in Section 2.2, a high number of the measurements were labeled as low quality, especially in the second scenario. This can be attributed to the low spatial or temporal resolution of the data, as we assume that the density of the entire pixel is the same as the average density at a ground measurement point. To overcome this, future studies may benefit from utilizing satellite imagery with higher spatial and temporal resolution, such as SuperDove data from the PlanetScope constellation.

Besides, as mentioned earlier, we utilized all bands and VIs in this study. This may be complex and increase the processing time for future cases, especially when the scale is regional and national. It will be even more computationally expensive if users want to use the time-series model, since it needs a total of 78 features. Therefore, feature engineering is of great importance for future work.

The trained models did not go through validation and verification outside of Canada or in eastern Canada. Collaboration with researchers within Canada and worldwide would be advantageous, provided they had reliable ground measurements to ensure the models work effectively across various study sites and crop circumstances.

5. Conclusion

The study illustrated the significant potential of Sentinel-2 and HLSL30 optical data and machine learning methods as a precise, practical, and cost-effective tool for automated stem counting and winter mortality detection. Two scenarios were assessed for training the machine learning models: 1) utilizing single-date remote sensing data and 2) employing time-series data. For the single-date stem count, the closest satellite data for the ground samples during the growing season were considered, while for the time-series data, three satellite measurements were utilized. Various ML classification and regression models have been trained over the ground-truth data. Classification models classified stem density into four categories (bare, low-density, medium-density, and high-density), achieving a maximum accuracy of 85 % with Sentinel-2 data and 84 % with HLSL30 data. The regression findings indicated that alfalfa stem density can be approximated with an error of approximately ± 6–9 stems/foot^2. The RF machine learning model demonstrated significant proficiency in estimating and classifying stem counts utilizing Sentinel-2 and HLSL30 datasets in both scenarios. Our findings also indicated that the NDVI and OSAVI were the most important features for alfalfa stem count analysis.

CRediT authorship contribution statement

Hazhir Bahrami: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Karem Chokmani:** Writing – review & editing, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Saeid Homayouni:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, Conceptualization. **Viacheslav I. Adamchuk:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Md Saifuzzaman:** Writing – review & editing, Methodology, Data curation. **Rami Albasha:** Writing – review & editing. **Maxime Leduc:** Supervision, Project administration, Funding acquisition.

Declaration of competing interest

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

Acknowledgment

We gratefully acknowledge the Agri-Risk program of Agriculture and Agri-Food Canada (grant number CASPP-040), Canadian Forage and Grassland Association (CFGA), Fonds de recherche du Québec - Nature et technologies (FRQNT), grant number 346438 (https://doi.org/10.69777/346438), Mitacs (grant number IT23205), De l'observation de la terre aux services d'information décisionnelle (DOTS), and the invaluable contributions of farmers and field advisors for their generous provision of data and financial support, without which this research would not have been possible. We also thank NVIDIA's Academic Grant program for providing two powerful NVIDIA Quadro RTX A6000 GPUs to our INRS Environmental and Northern Remote Sensing Laboratory.

Data availability

Due to the data sharing policy outlined in our contract with our partners, we are currently (July 04, 2025) unable to release the in-situ measurements and trained models.

References

Atzberger, C., 2013. Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs. Remote Sens. 5, 949–981.

Awad, M., Khanna, R., Awad, M., Khanna, R., 2015. Support vector regression. Efficient learning machines: Theories, concepts, and applications for engineers and system designers, 67-80.

Bahrami, H., Chokmani, K., Homayouni, S., Adamchuk, V.I., Albasha, R., Saifuzzaman, M., Leduc, M., 2025a. Machine learning-based alfalfa height estimation using sentinel-2 multispectral imagery. Remote Sens. 17, 1759.

Bahrami, H., Chokmani, K., Homayouni, S., Adamchuk, V.I., Saifuzzaman, M., Leduc, M., 2025b. Alfalfa detection and stem count from proximal images using a combination of deep neural networks and machine learning. Comput. Electron. Agric. 232, 110115.

Bahrami, H., Homayouni, S., McNairn, H., Hosseini, M., Mahdianpari, M., 2022a. Regional crop characterization using multi-temporal optical and synthetic aperture radar earth observations data. Can. J. Remote Sens. 48, 258–277.

Bahrami, H., Homayouni, S., Safari, A., Mirzaei, S., Mahdianpari, M., Reisi-Gahrouei, O., 2021. Deep learning-based estimation of crop biophysical parameters using multi-source and multi-temporal remote sensing observations. Agron 11, 1363.

Bahrami, H., McNairn, H., Mahdianpari, M., Homayouni, S., 2022b. A meta-analysis of remote sensing technologies and methodologies for crop characterization. Remote Sens. 14, 5633.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brownlee, J., 2016. XGBoost with python: gradient boosted trees with XGBoost and scikit-learn. Machine Learning Mastery.

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series landsat data and a machine learning approach. Remote Sens. Environ. 210, 35–47.

Canada, S., 2021. Table 32-10-0249-01 Land use, Census of Agriculture.

Castonguay, Y., Laberge, S., Brummer, E.C., Volenec, J.J., 2006. Alfalfa winter hardiness: a research retrospective and integrated perspective. Adv. Agron. 90, 203–265.

Chandel, A.K., Khot, L.R., Yu, L.-X., 2021. Alfalfa (Medicago sativa L.) crop vigor and yield characterization using high-resolution aerial multispectral and thermal infrared imaging technique. Comput. Electron Agric. 182, 105999.

Chen, J., Yu, T., Cherney, J.H., Zhang, Z., 2024. Optimal integration of optical and SAR data for improving alfalfa yield and quality traits prediction: new insights into satellite-based forage crop monitoring. Remote Sens. 16, 734.

Cherney, J.H., Kallenbach, R.L., Picasso Risso, V.D., 2020. Systems for temperate humid areas. Forages: Sci. Grassland Agric. 2, 369–386.

Chlingaryan, A., Sukkarieh, S., Whelan, B., 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. Comput. Electron Agric. 151, 61–69.

Claverie, M., Masek, J.G., Ju, J., Dungan, J.L., 2017. Harmonized landsat-8 sentinel-2 (HLS) product user's guide. National Aeronautics and Space Administration (NASA): Washington, DC, USA.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297.

Dan Undersander, C.G., Dennis Cosgrove, Jerry Doll, and Neal Martin, 2011. Alfalfa Stand Assessment: Is This Stand Good Enough to Keep?.

Dangeti, P., 2017. Statistics for machine learning. Packt Publishing Ltd.

Daughtry, C.S., Walthall, C., Kim, M., De Colstoun, E.B., McMurtrey Iii, J., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sens. Environ. 74, 229–239.

Delegido, J., Verrelst, J., Alonso, L., Moreno, J., 2011. Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. Sens 11, 7063–7081.

Dong, T., Liu, J., Qian, B., He, L., Liu, J., Wang, R., Jing, Q., Champagne, C., McNairn, H., Powers, J., 2020. Estimating crop biomass using leaf area index derived from Landsat 8 and Sentinel-2 data. ISPRS J. Photogramm. Remote Sens. 168, 236–250.

Dvorak, J.S., Pampolini, L.F., Jackson, J.J., Seyyedhasani, H., Sama, M.P., Goff, B., 2021. Predicting quality and yield of growing alfalfa from a UAV. Trans. ASABE 64, 63–72.

Echeverría, A., Urmeneta, A., González-Audícana, M., González, E.M., 2021. Monitoring rainfed alfalfa growth in semiarid agrosystems using Sentinel-2 imagery. Remote Sens. 13, 4719.

Falanga Bolognesi, S., Pasolli, E., Belfiore, O.R., De Michele, C., D'Urso, G., 2020. Harmonized landsat 8 and sentinel-2 time series data to detect irrigated areas: an application in Southern Italy. Remote Sens. 12, 1275.

Feng, L., Zhang, Z., Ma, Y., Du, Q., Williams, P., Drewry, J., Luck, B., 2020. Alfalfa yield prediction using UAV-based hyperspectral imagery and ensemble learning. Remote Sens. 12, 2028.

Frampton, W.J., Dash, J., Watmough, G., Milton, E.J., 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS J. Photogramm. Remote Sens. 82, 83–92.

Gamon, J., Surfus, J., 1999. Assessing leaf pigment content and activity with a reflectometer. The New Phytol. 143, 105–117.

Gao, S., Zhong, R., Yan, K., Ma, X., Chen, X., Pu, J., Gao, S., Qi, J., Yin, G., Myneni, R.B., 2023. Evaluating the saturation effect of vegetation indices in forests using 3D radiative transfer simulations and satellite observations. Remote Sens. Environ. 295, 113665.

Giridhar, K., Samireddypalle, A., 2015. Impact of climate change on forage availability for livestock. Climate Change Impact on Livestock: Adaptation and Mitigation, 97-112.

Gitelson, A., Merzlyak, M.N., 1994. Spectral reflectance changes associated with autumn senescence of Aesculus hippocastanum L. and Acer platanoides L. leaves. Spectral features and relation to chlorophyll estimation. J. Plant Physiol. 143, 286–292.
Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. Remote Sens. Environ. 80, 76–87.
Gitelson, A.A., Viña, A., Arkebauer, T.J., Rundquist, D.C., Keydan, G., Leavitt, B., 2003. Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophys. Res. Lett. 30.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27.
Huete, A., Liu, H., Batchily, K., Van Leeuwen, W., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sens. Environ. 59, 440–451.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.
Jordan, C.F., 1969. Derivation of leaf-area index from quality of light on the forest floor. Ecology 50, 663–666.
Kamilaris, A., Kartakoullis, A., Prenafeta-Boldú, F.X., 2017. A review on the practice of big data analysis in agriculture. Comput. Electron Agric. 143, 23–37.
Kayad, A.G., Al-Gaadi, K.A., Tola, E., Madugundu, R., Zeyada, A.M., Kalaitzidis, C., 2016. Assessing the spatial variability of alfalfa yield using satellite imagery and ground-based data. PLoS One 11, e0157166.
Kganyago, M., Adjorlolo, C., Mhangara, P., Tsoeleng, L., 2024. Optical remote sensing of crop biophysical and biochemical parameters: an overview of advances in sensor technologies and machine learning algorithms for precision agriculture. Comput. Electron Agric. 218, 108730.
Kganyago, M., Mhangara, P., Adjorlolo, C., 2021. Estimating crop biophysical parameters using machine learning algorithms and Sentinel-2 imagery. Remote Sens. 13, 4314.
Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine learning in agriculture: a review. Sens 18, 2674.
Maxime Leduc, R.B., 2019. Did your alfalfa stand survive winter?, p. 18.
McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432.
McKenzie, J., Paquin, R., Duke, S.H., 1988. Cold and heat tolerance. Alfalfa and Alfalfa Improv. 29, 259–302.
Meroni, M., Marinho, E., Sghaier, N., Verstrate, M.M., Leo, O., 2013. Remote sensing based yield estimation in a stochastic framework—case study of durum wheat in Tunisia. Remote Sens. 5, 539–557.
Noland, R.L., Wells, M.S., Coulter, J.A., Tiede, T., Baker, J.M., Martinson, K.L., Sheaffer, C.C., 2018. Estimating alfalfa yield and nutritive value using remote sensing and air temperature. Field Crop Res 222, 189–196.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.
Pokhariyal, S., Patel, N., Govind, A., 2023. Machine learning-driven remote sensing applications for agriculture in India—a systematic review. Agron 13, 2302.
Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sens. Environ. 48, 119–126.
Qiao, K., Zhu, W., Xie, Z., Wu, S., Li, S., 2024. New three red-edge vegetation index (VI3RE) for crop seasonal LAI prediction using Sentinel-2 data. Int. J. Appl. Earth Obs. Geoinf. 130, 103894.
Rezaei, E.E., Webber, H., Asseng, S., Boote, K., Durand, J.L., Ewert, F., Martre, P., MacCarthy, D.S., 2023. Climate change impacts on crop yields. Nat. Rev. Earth Environ. 4, 831–846.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. NASA Spec. Publ. 351, 309.
Saifuzzaman, M., Adamchuk, V., Leduc, M., 2022. Stem Characteristics and Local Environmental Variables for Assessment of Alfalfa Winter Survival.
Sapkota, A., Haghverdi, A., Montazar, A., 2023. Estimating fall-harvested alfalfa (Medicago sativa L.) yield using unmanned aerial vehicle–based multispectral and thermal images in southern California. Agrosystems, Geosci. Environ. 6, e20392.
Sathyanarayanan, S., Tantri, B.R., 2024. Confusion matrix-based performance evaluation metrics. Afr. J. Biomed. Res. 27, 4023–4031.
Schwalbert, R.A., Amado, T.J., Nieto, L., Varela, S., Corassa, G.M., Horbe, T.A., Rice, C. W., Peralta, N.R., Ciampitti, I.A., 2018. Forecasting maize yield at field scale based on high-resolution satellite imagery. Biosystems Eng. 171, 179–192.
Tedesco, D., Nieto, L., Hernández, C., Rybecky, J.F., Min, D., Sharda, A., Hamilton, K.J., Ciampitti, I.A., 2022. Remote sensing on alfalfa as an approach to optimize production outcomes: a review of evidence and directions for future assessments. Remote Sens. 14, 4940.
Thivierge, M.N., Bélanger, G., Jégo, G., Delmotte, S., Rotz, C.A., Charbonneau, É., 2023. Perennial forages in cold-humid areas: Adaptation and resilience-building strategies toward climate change. Agron. J. 115, 1519–1542.
Undersander, D., Grau, C., Cosgrove, D., Doll, J., Martin, N., 1998. Alfalfa stand assessment: is this stand good enough to keep? University of Wisconsin-Extension, Madison, WI, p. 3620.
Van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: a systematic literature review. Comput. Electron Agric. 177, 105709.
Vincini, M., Frazzi, E., 2011. Comparing narrow and broad-band vegetation indices to estimate leaf chlorophyll content in planophile crop canopies. Precis. Agric. 12, 334–344.
Wang, X., Cai, G., Lu, X., Yang, Z., Zhang, X., Zhang, Q., 2022. Inversion of wheat leaf area index by multivariate red-edge spectral vegetation index. Sustain 14, 15875.
Wang, X., Yu, S., Wen, Z., Zhang, L., Fang, C., Jiang, L., 2023. Application of modern GIS and remote sensing technology based on big data analysis in intelligent agriculture. J. Indian Soc. Remote Sens. 51, 1891–1901.
Xiaoqin, W., Miaomiao, W., Shaoqiang, W., Yundong, W., 2015. Extraction of vegetation information from visible unmanned aerial vehicle images. Trans. Chin. Soc. Agri. Eng. 31.
Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. J. Plant Ecol. 1, 9–23.
Xu, J., Du, X., Dong, T., Li, Q., Zhang, Y., Wang, H., Liu, M., Zhu, J., Yang, J., 2025. Estimation of sugarcane biomass from Sentinel-2 leaf area index using an improved SAFY model (SAFY-Sugar). Int. J. Appl. Earth Obs. Geoinf. 140, 104570.
Zhang, P.-P., Zhou, X.-X., Wang, Z.-X., Mao, W., Li, W.-X., Yun, F., Guo, W.-S., Tan, C.-W., 2020. Using HJ-CCD image and PLS algorithm to estimate the yield of field-grown winter wheat. Sci. Reports 10, 5173.
Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D.B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., 2017. Temperature increase reduces global yields of major crops in four independent estimates. Proc. Nat. Acad. Sci. 114, 9326–9331.