

## Unmanned aerial vehicle (UAV) imaging and machine learning applications for plant phenotyping

Fitsum T Teshome <sup>a</sup>, Haimanote K Bayabil <sup>a,\*</sup>, Gerrit Hoogenboom <sup>b,c</sup>, Bruce Schaffer <sup>d</sup>, Aditya Singh <sup>b</sup>, Yiannis Ampatzidis <sup>e</sup>

<sup>a</sup> Department of Agricultural and Biological Engineering, Tropical Research and Education Center, IFAS, University of Florida, Homestead, FL 33031, USA

<sup>b</sup> Department of Agricultural and Biological Engineering and Institute for Sustainable Food Systems, University of Florida, Gainesville, FL 32603, USA

<sup>c</sup> Global Food Systems Institute, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL 32611-0910, USA

<sup>d</sup> Horticultural Sciences Department, Tropical Research and Education Center, IFAS, University of Florida, Homestead, FL 33031, USA

<sup>e</sup> Department of Agricultural and Biological Engineering, Southwest Florida Research and Education Center, University of Florida, IFAS, Immokalee, FL 34142, USA



### ARTICLE INFO

#### Keywords:

Crop surface model  
K-nearest neighbor  
Plant biomass  
Plant height, random forest  
Support vector machine

### ABSTRACT

The availability of timely and accurate information about plant conditions is critical for making informed decisions for maintaining or improving agricultural productivity. However, field data collection is often laborious and time-consuming, underscoring the need for accurate and efficient methods to monitor crop growth and development. Therefore, the objective of this study was to evaluate the effectiveness of combining unmanned aerial vehicle (UAV)-based imaging and machine learning (ML) techniques for monitoring sweet corn (*Zea mays* var. *saccharata*) height, biomass, and yield. The study was conducted at the Tropical Research and Education Center (TREC) during the winter (dry) season of 2020–2021 using 16 experimental plots planted with sweet corn. The treatments were set up in a completely randomized block design (RCBD) with four irrigation treatments of 25%, 50%, 75%, and 100% full irrigation with four replications each. Field data collection included plant height, fresh and dry biomass, and yield. In addition, UAV images were collected using the DJI Matrice 210 v2 UAV (SZ DJI Technology Co., Ltd., Shenzhen, China) equipped with a MicaSense RedEdge-MX multispectral sensor (MicaSense, Seattle, WA, USA). Image processing was done with Pix4Dmapper 4.7.5 (Pix4D S.A., Prilly, Switzerland). A crop surface model, representing estimated plant height (UAVH), was calculated based on pixel-to-pixel differences between digital surface and terrain models. A simple linear regression model was used to estimate sweet corn biomass and yield from UAV images estimated plant height (UAVH). In addition, two linear algorithms known as a linear model (LM) and Lasso and elastic-net regularized generalized linear model (GLMNET) and three non-linear ML algorithms including random forest (RF), support vector machine (SVM), and k-nearest neighbor (kNN) were used to predict plant height and biomass. These algorithms were chosen due to their reliable performance and ability to learn complex non-linear relationships. Eight vegetation indices with UAVH were also used to evaluate the ML models' performance to predict plant height and biomass. Results confirmed that UAV imaging could be effectively used to estimate plant height ( $d = 0.99$ ,  $r^2 = 0.99$ , and MAE = 5 cm). Estimated fresh and dry biomass from UAV imaging also had good agreements with measured data with  $r^2$  values of 0.75 and 0.70, respectively. A statistically significant linear correlation between measured fresh yield at harvest and UAVH was found with  $d$ , adjusted  $r^2$ , and MAE of 0.84, 0.66, and 67 g m<sup>-2</sup>, respectively. Evaluation of ML algorithms revealed that all models performed well for plant biomass estimation with  $d$  values between 0.88 and 0.99. However, the kNN and SVM outperformed all other models for biomass estimation. GLMNET performed better than other models for plant height estimation. Overall, results revealed that UAV imaging and ML models could be effectively used for monitoring plant phenotypic characteristics such as height, yield, and biomass.

\* Corresponding author.

E-mail address: [hbayabil@ufl.edu](mailto:hbayabil@ufl.edu) (H.K. Bayabil).

## 1. Introduction

The lack of efficient crop monitoring practices is one of the factors that have contributed to the decline of agricultural production (Becker-Reshef et al., 2020; Esquerdo et al., 2011; Kalogiannidis et al., 2022; Löw et al., 2018; Wiseman et al., 2014). Agricultural producers have been able to improve their efficiency and effectiveness in monitoring and controlling various aspects of their operations using remote sensing imageries (Barrientos et al., 2011; Delavarpour et al., 2021; Huang et al., 2018b, 2018a; Paul J. Pinter et al., 2003). This includes the detection of plant populations (Ampatzidis et al., 2020; Ampatzidis and Partel, 2019), plant biomass and yield (Battude et al., 2016; Johnson, 2013; Lobell et al., 2003; Prasad et al., 2006), insect damage (Adelabu et al., 2014; Marston et al., 2020; Oumar and Mutanga, 2014), herbicide damage (Dicke et al., 2012; Robles et al., 2010; Thelen et al., 2004), water (Ismail and Mutanga, 2010; Zhang et al., 2019), and nutrient deficiencies (Candiago et al., 2015), and weed infestations (Rasmussen et al., 2016; Samseemoung et al., 2012; Goel et al., 2003; Huang et al., 2018b). Coarse spatial and temporal resolutions of satellite images are always a challenge for agriculture applications (Ahamed et al., 2011; Jung et al., 2021; Moran et al., 1997; Veysi et al., 2017; Zhang and Kovacs, 2012). In addition, satellite image processing is a very laborious and time-consuming process (Guo et al., 2019; Rajaveni et al., 2017; Xu and Ghamisi, 2022). However, the autonomous flight capability (Klemas, 2015), spatial and temporal resolution flexibility (Doughty and Cavanaugh, 2019), and cost-effectiveness (Singh and Frazier, 2018) have given unmanned aerial vehicle (UAV) more popularity (Ćwiąkala et al., 2018; Easterday et al., 2019; Pérez-Ortiz et al., 2015; Zhang and Kovacs, 2012). UAVs have been used commercially since the early 1980s (Crampton, 2016; Huang et al., 2013; Lattanzi and Miller, 2017; Mazur, 2016; Rosser et al., 2018). Since then, the practical applications of UAVs have skyrocketed in various industries. UAVs have been effectively employed for broad applications such as agriculture and forest monitoring, disaster management, traffic surveillance, and photogrammetry for 3D modeling (Gupta et al., 2021; Li and Liu, 2019; Nex and Remondino, 2014; Yao et al., 2019; Yu et al., 2010).

UAV images have shown promising potential for improving agricultural production (Huang et al., 2013; Marston et al., 2020). UAV images could efficiently replace traditional crop scouting and phenotyping, which is often a laborious, time-consuming, and a subjective process (Atefi et al., 2021; Deery and Jones, 2021; Partel et al., 2019; Zhang et al., 2022). UAV-based imaging improves the accuracy and efficiency of phenotyping by regularly monitoring crops' status without plant or soil destruction (Holman et al., 2016). Several researchers have considered fine-resolution UAV images for crop phenotyping parameters, such as plant height, biomass, and yield measurements (Poudyal et al., 2023; Vijayakumar et al., 2023). Han et al. (2018) found strong correlations between UAV images estimated maize (*Zea mays L.*) height and ground truth height. Watanabe et al. (2017) found that UAV-based remote sensing could improve the efficiency of a high-throughput phenotyping system for sorghum (*Sorghum bicolor (L.) Moench*) plant height and its response to nitrogen availability. Anthony et al. (2014) revealed that despite the high variability of UAV-mounted laser scanners, they were able to precisely fly over the crops and determine the height of crops in a corn (*Zea mays*) field and indoor testbed. UAV-mounted remote sensing sensors have great potential for crop phenotyping and yield estimation; however, the increasing number of data being collected through UAV-based remote sensing techniques has made it difficult to easily understand the relationship between plant phenotypic parameters such as plant height, plant biomass, vegetation indices, and yield (Selvaraj et al., 2020; Reedha et al., 2022). Conventional models, such as the simple linear regression model, could fail to identify the relationship between the parameters (Mochida et al., 2019; Shakoor et al., 2017).

An artificial intelligence (AI) algorithm can learn to identify the relationship between dependent variables, such as yield and biomass,

and various independent plant traits, such as plant height and vegetation indices, using machine learning (ML) algorithms (Delen et al., 2010; Guo et al., 2020; van Klompenburg et al., 2020; Wang et al., 2016). Numerous researchers have shown artificial intelligence models, i.e., machine learning and deep learning models, could be potential tools for high throughput plant phenotyping and disease detection (C. K. et al., 2022; C.k. et al., 2023; Mochida et al., 2019; Pound et al., 2017; Ranparia et al., 2020; Sunil et al., 2022; Ubbens and Stavness, 2017).

Regression is a process utilized in machine learning to identify the patterns and relationships in a dataset. It can then be used to predict the outcome of new data or fill in the gap left by missing information (Alamoodi et al., 2021; Lakshminarayan et al., 1999; Wang et al., 2022). Linear regression (LM) and Lasso and elastic-net regularized generalized linear models (GLMNEN) are linear algorithms commonly used in supervised learning (Jiménez et al., 2019; Khanal et al., 2018; Kopitar et al., 2020; Ren et al., 2020; Sun et al., 2020). Support vector machines (SVM), k-nearest neighbors (kNN), and random forest (RF) are the most common non-linear ML algorithms used to identify the relationships within the data (Alonso et al., 2018; Mohammed et al., 2019; Rashidi et al., 2019). The ML algorithms could improve and fasten the process of detecting the relationships between plant phenotypic parameters, i.e., plant height, biomass, and yield (Linaza et al., 2021; Soltis et al., 2020; Yalcin, 2017).

UAV imaging coupled with machine learning models provides a state-of-the-art tool for high-throughput plant phenotyping studies and offers more detailed, rapid, non-destructive data collection and automated analysis capabilities. This helps plant breeders and crop managers to understand plant growth and development, aids in crop production improvement, precision agriculture, and investigates plant responses to different management conditions.

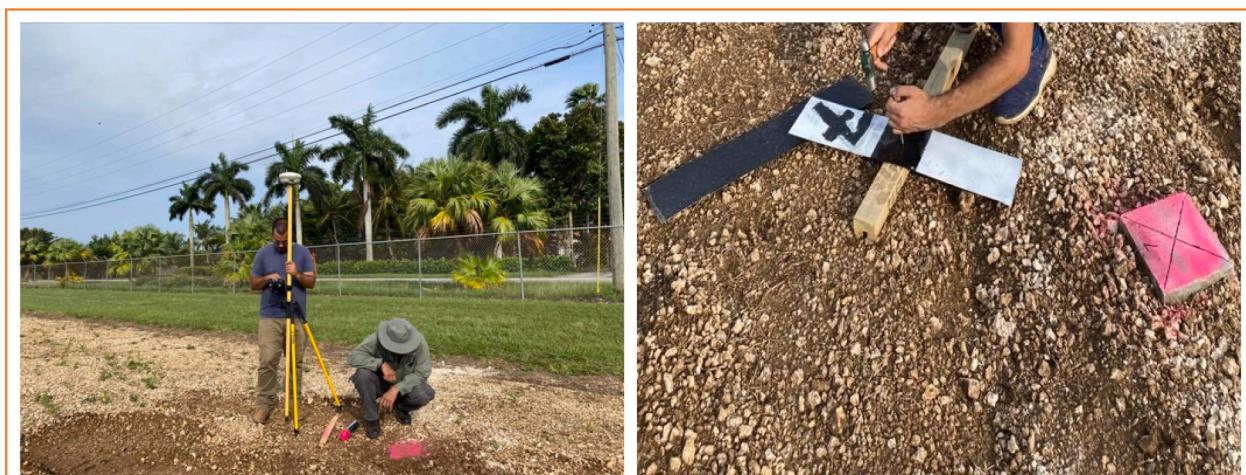
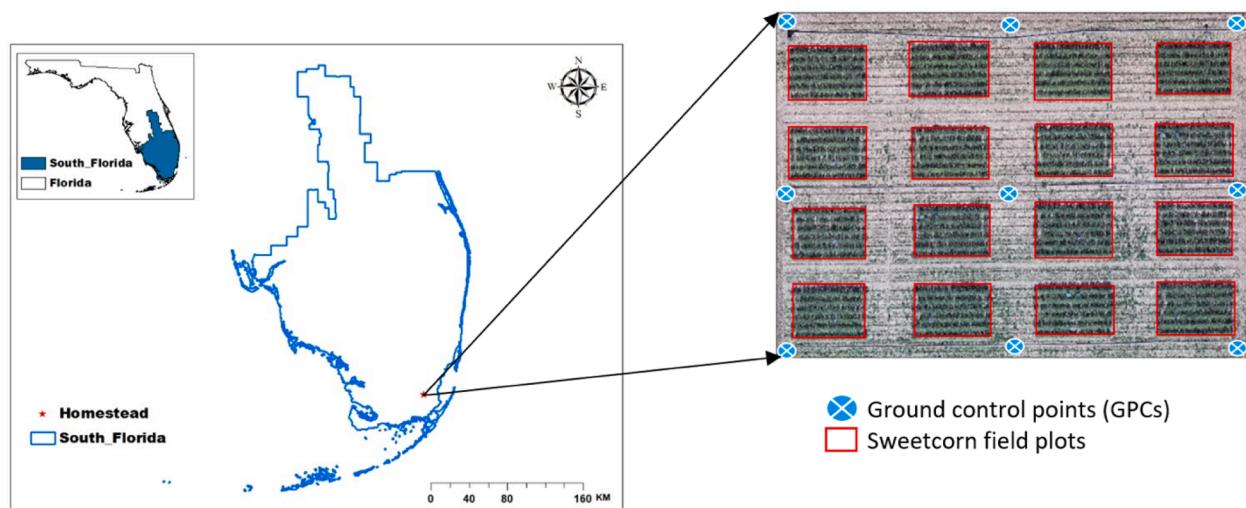
The development of a high-throughput phenotyping system coupled with machine learning models for the production of sweet corn could help boost the vegetable's production in Florida, which is already known as one of the country's top sweet corn-producing states (Dittmar et al., 2021). Therefore, this study was conducted to evaluate the effectiveness of combining unmanned aerial vehicle (UAV)-based imaging and machine learning (ML) techniques for monitoring sweet corn (*Zea mays var. saccharata*) height, biomass, and yield, with the aim of providing a more accurate and efficient means of monitoring crop growth and development.

## 2. Materials and methods

### 2.1. Research site

This experiment was conducted at the University of Florida's Tropical Research and Education Center (TREC) in Homestead, Florida, USA (Fig. 1). The research site is located at 25.51° N and to 80.50° W. The experimental plots covered an area of approximately 5,666 m<sup>2</sup> (Fig. 1). Sweet corn (Battalion cultivar) was grown on 16 plots (9 m × 5.5 m) from 24 Nov 2020 to 19 Feb 2021. Each plot had six rows with 91 cm spacing. The plots were planted using a row crop planter with 8–13 cm plant spacing. The treatments were set up in a completely randomized block design (RCBD) with four irrigation treatments of 25, 50, 75, and 100% maximum allowable depletion (MAD) with four replications each. The plots were established under a linear move sprinkler irrigation system.

A total of nine ground control points (GCPs) were marked with 20 x 51 cm concrete blocks (Fig. 2b) and uniformly installed around and within the experimental field (Fig. 1). The geographic coordinates of each concrete block center were measured using a global navigation satellite system (GNSS) real-time kinematic (RTK) system (Trimble R4 GNSS, Trimble, Sunnyvale, CA, USA) (Fig. 2a). White and black (15 x 76 cm) painted aluminum sheets were drilled in the top of the concrete blocks so that they could be identified from UAV images (Fig. 2b). Five out of the nine GCPs were used to geo-reference the UAV ortho-mosaic



**Fig. 2.** Installing ground control points (GCPs). Measuring geographic coordinates of GCPs using a global navigation satellite system (GNSS) real-time kinematic (RTK) (a) and drilling black and white painted aluminum sheets on concrete blocks (b).



**Fig. 3.** Ground height measurements using a measuring tape (a and b) and recording height (c).

images to minimize geometric distortions. The geo-referencing process was done using Pix4Dmapper 4.7.5 (Pix4D S.A., Prilly, Switzerland). The remaining four points were used as checkpoints (CPs) to validate the accuracy of the geo-referencing.

## 2.2. Ground measurements of plant phenotypic characteristics

Ground measurements included plant height, above-ground biomass, and, at the end of the experiment, cob yield. Plant height and biomass measurements were collected bi-weekly to capture the different crop growth stages. From each plot, four plants were identified from the inside rows for height measurements. Using a measuring tape, plant height was measured from the soil surface to the plant apex (Fig. 3). Five plant height measurements (3 Feb 2021, 20 Jan 2021, 6 Jan 2020, 23 Dec 2020, and 9 Dec 2020) were recorded during the crop-growing season. Four biomass measurements (3 Feb 2021, 20 Jan 2021, 6 Jan 2020, and 23 Dec 2020) were conducted, which were harvested from a 30 cm length row that consisted of three plants on average. Outside rows were avoided to minimize border effects. Each plant was cut 13 mm above the ground surface. After each harvest, the total fresh biomass of the plants was immediately determined. The fresh leaf and stem were then separated and weighed. The total dry biomass was measured after the harvested plants were oven dried for 48 h at 80 °C. For the final cob yield and biomass measurements (22 Feb 2021), the plants were harvested from a set of 6 m length rows. The cobs were separated from the vegetative plant parts and weighed. Biomass per unit area was estimated by dividing the total harvested plant biomass by the harvested land area.

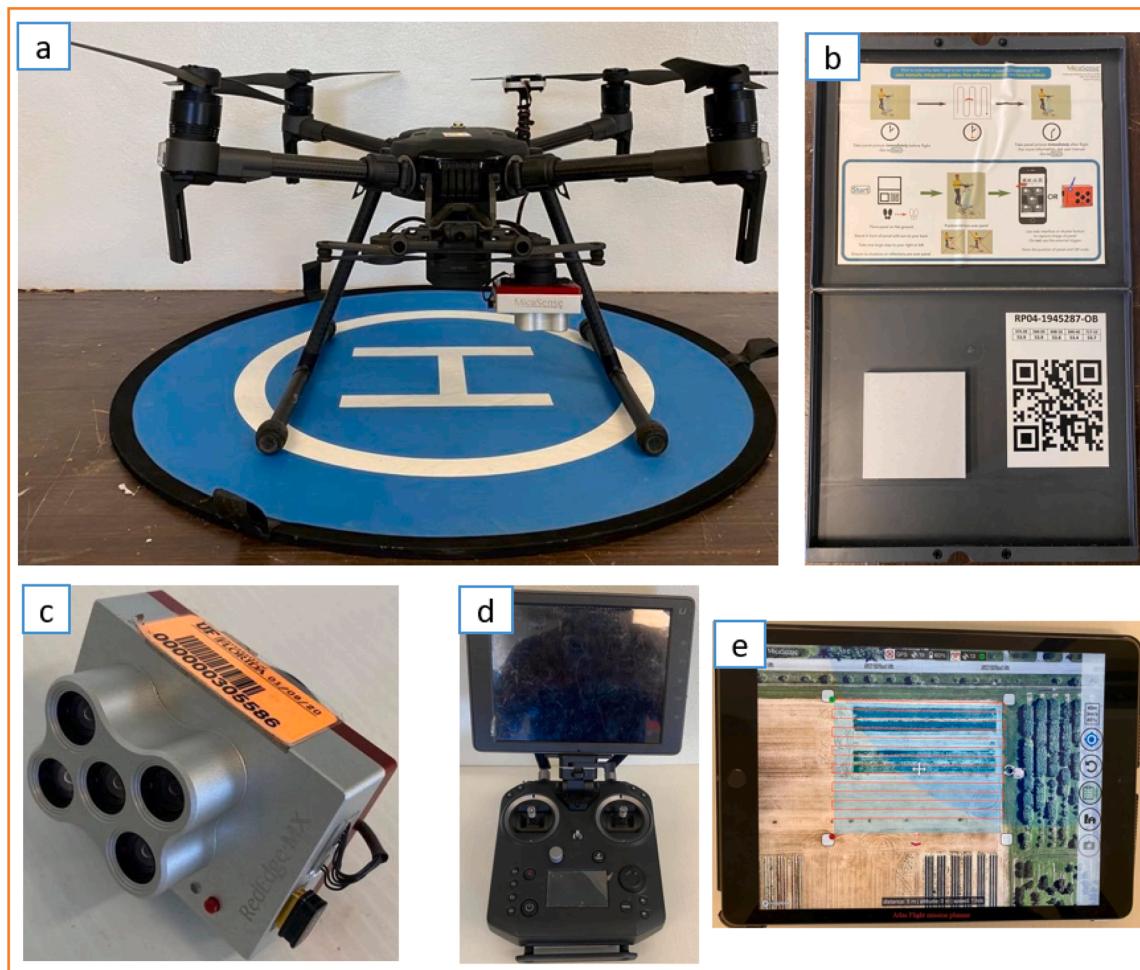
The same procedure was followed to estimate yield per unit area.

## 2.3. UAV images collection and processing

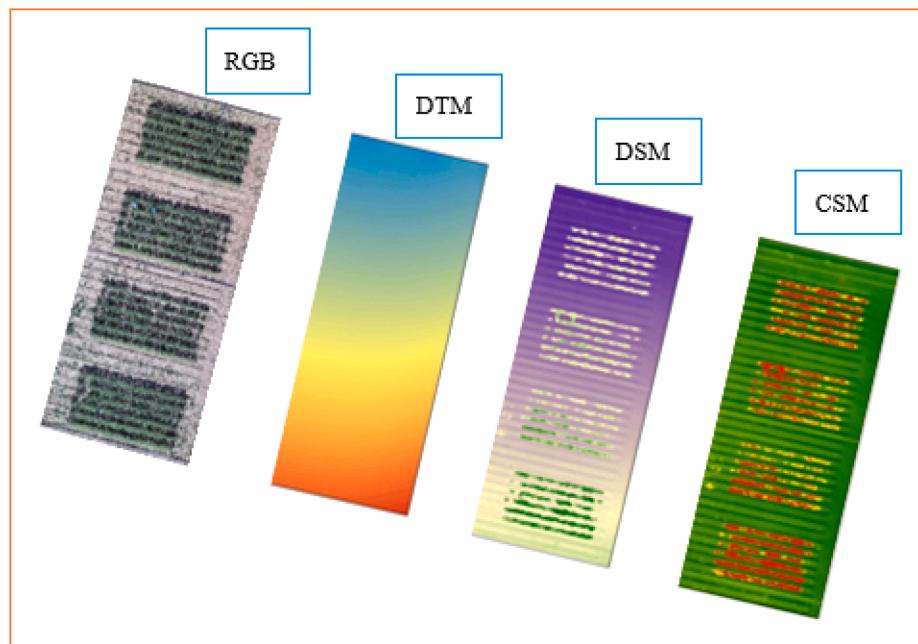
High-resolution multispectral imageries were collected daily during the entire growing season using a RedEdge-MX sensor (MicaSense, Seattle, WA, USA). The sensor has five spectral bands, Red (center wavelength, 668 nm), Green (560 nm), Blue (475 nm), Near-Infrared (842 nm), and RedEdge (717 nm). The Atlas Flight mission planning application (MicaSense, 2021) was used to generate predefined autonomous flight missions and fly the UAV (Fig. 4). The planner was programmed to autonomously fly DJI Matrices 210 v2 UAV (SZ DJI Technology Co., Ltd., Shenzhen, China) at 40 m above ground level (AGL) with a flight speed of 3 m s<sup>-1</sup>. Images were collected daily at solar noon with 85% forward and side overlaps. Approximately, 2,500 images were collected per flight with an average ground sampling distance (GSD) of 1.03 cm pixel<sup>-1</sup>. Before and after each flight, images of a calibration panel (Fig. 4b) were manually collected using the RedEdge-MX camera. The images from the calibration panel were then used to calibrate and correct images collected on the same day using Pix4Dmapper 4.7.5.

## 2.4. Development of digital surface and terrain models

The digital surface model (DSM) shows the elevation of any feature above mean sea level, whereas the digital terrain model (DTM) represents continuous ground surface elevation where all non-terrain objects



**Fig. 4.** DJI Matrices 210 v2 UAV (a), MicaSense calibration reflectance panel (b), Rededge-MX camera (c), DJI remote controller (d), and Atlas Flight mission planner (iOS version) (e).



**Fig. 5.** Red-green-blue (RGB), digital terrain model (DTM), digital surface model (DSM), and crop surface model (CSM) of four plots of a sweet corn field. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

were removed (Fig. 5). Pix4Dmapper 4.7.5 was used to generate the DSM and DTM. The pixel level difference between DSM and DTM was estimated using ArcGIS 10.7.1 (ESRI, Redlands, CA, USA), to obtain the crop surface model (CSM) (Fig. 5).

## 2.5. Estimating plant height using UAV images

Geographic coordinates of ground height measurements were used to extract estimated height from the CSMs, which are referred to as UAV images estimated plant height (hereafter referred to as UAVH). The UAVH was then compared with ground height measurements (Fig. 6). The coefficient of determination ( $r^2$ ), index of agreement (d), mean absolute error (MAE), and the root mean square error (RMSE) were used to evaluate the agreement between measured and estimated heights.

## 2.6. Estimating plant biomass and yield using simple linear regression

Field-based measurements (total fresh biomass, fresh leaf biomass, fresh stem biomass, and total dry biomass) were regressed against UAV-estimated plant heights utilizing simple linear regressions (Fig. 6). To avoid introducing a bias into the model evaluation, the data were randomly separated into validation and evaluation periods. Out of the total data collected during the experiment about 60% of the data were used to develop a regression model with the remaining 40% of the data held back for cross-validation.

Partitioning data into 30/70% for simple linear regression models is not recommended because it can lead to overfitting (Cosenza et al., 2022). In a simple linear regression model, the goal is to fit a straight line through the data that represents the relationship between the independent variable and the dependent variable (Zou et al., 2003). If a large portion of the data is used for training, the model may become too complex and overfit the training data, resulting in poor performance when applied to new data. Therefore, it is recommended to use cross-validation techniques to assess model performance and prevent overfitting (Berrar, 2018). Cross-validation involves dividing the data into smaller subsets and testing the model on each subset, which helps to assess the model's performance on unseen data.

### 2.6.1. A simple linear regression model

Simple linear regression is a type of statistical method that can be used to establish a relationship between two variables by fitting a straight line to the data. This method is usually performed on the assumption that the relationship is linear and that the errors are normally distributed. A simple linear regression model only involves one dependent and one independent variable.

## 2.7. Estimating plant phenotype using machine learning models

### 2.7.1. Vegetation indices

We derived eight vegetation indices (VIs) from UAV imagery (Table 1). The selection of VIs was based on a literature review of previous studies that reported these indices could be reliably used to estimate plant height and biomass. The UAV images estimated vegetation indices and plant height (UAVH) were used to evaluate the performance of five machine learning models (SVM, kNN, RF, LM, and GLMNET) for estimating sweet corn biomass. However, only VIs were used to evaluate the performance of ML models for sweet corn height estimation.

### 2.7.2. Machine learning models

Regression algorithms in ML are a type of supervised learning algorithm that requires labeled data to train and test a model (Mahesh, 2019). The ML algorithms were tuned to provide better performance in comparison to simple linear regression models proposed for estimating biomass and yields using UAV images. In this study, two linear algorithms known as LM and GLMNET and three non-linear machine learning algorithms, including RF, SVM, and kNN were used to predict plant height and biomass (Fig. 7). The datasets used for the evaluation of machine learning models for plant height, biomass, and yield estimation were randomly partitioned into 70% training data and 30% testing data. Furthermore, k-fold cross-validation ( $k = \text{five}$ ) was used to separate training data into training and validation datasets to indicate the generalization performance of the model across the entire dataset.

**2.7.2.1. Linear regression.** Linear regression (LM) is used in predictive analysis. It predicts the outcomes of various types of variables by showing a linear relationship between one or more independent and dependent elements (Osborne and Waters, 2019). It can then predict the

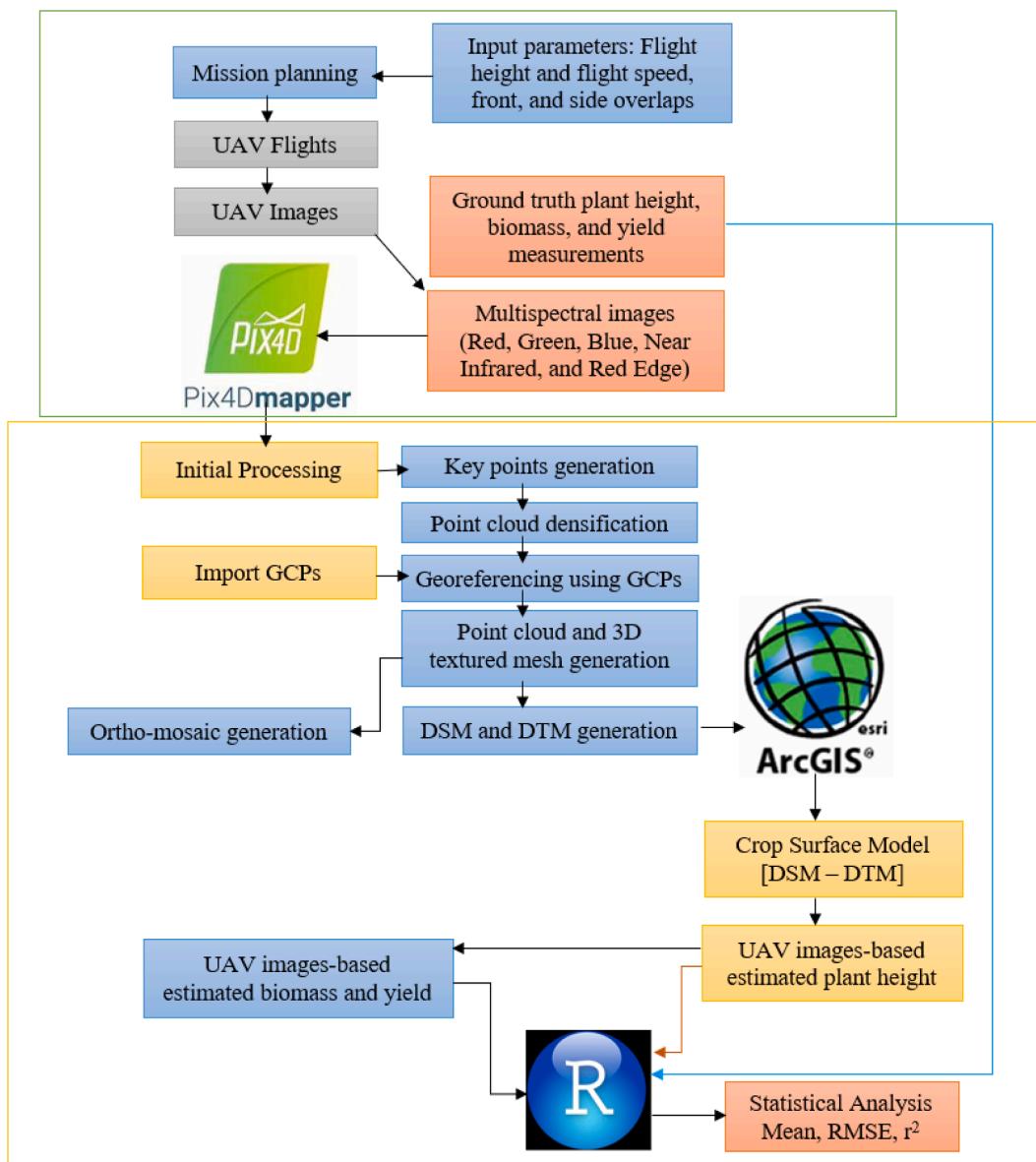


Fig. 6. General workflow for the estimation of plant height, yield, and biomass for sweet corn from UAV imagery.

outcome of the variable based on its relationship with the independent variable (Maulud and Abdulazeez, 2020).

**2.7.2.2. Lasso and elastic-net regularized generalized linear model.** Lasso and elastic-net regularized generalized linear model (GLMNET) is designed for fitting generalized linear models with a penalty maximum likelihood. It considers the elastic net or lasso penalty and computes a regularization path (de Vlaming and Groenen, 2015). The algorithms use a cyclical coordinate descent to calculate timings along a regularization path and they can handle sparse features and large problems (Friedman et al., 2010).

**2.7.2.3. Support vector machine.** The support vector machine (SVM) is a powerful tool for solving classification and regression problems due to its ability to focus on complex input variables' dynamics (Cortes and Vapnik, 1995). It utilizes a set of kernel functions, such as the linear, polynomial, and radial basis functions. These functions can transform the low-dimensional input space into an increased dimensional space (Yu et al., 2010).

**2.7.2.4. K-nearest neighbor.** The k-nearest neighbor (kNN) algorithm is a lazy and non-parametric supervised ML algorithm (Ali et al., 2019). This method works by plotting each point in a high-dimensional space, with each axis representing an individual variable (Nadkarni, 2016). The main disadvantage of the kNN method is that it is affected by redundant features. However, it is very effective and easy to implement (Nadkarni, 2016).

**2.7.2.5. Random forest.** Random forest (RF) is a supervised learning algorithm that could be used for regression and classification problems (Boulesteix et al., 2012). It works by randomly selecting a training sample to create independent decision trees (Svetnik et al., 2003). After getting the predicted results from the various decision trees, the final results of the analysis are determined by averaging the results for each of the problems for regression and taking their majority vote for the classification (Strobl et al., 2009).

#### 2.7.3. Feature selection

We employed best subsets regression analysis to select important input features for predicting plant phenotypic parameters

**Table 1**

List of vegetation indices used for plant height and biomass estimation with the machine learning models.

No	Vegetation indices	Equations	References
1	2-band Enhanced Vegetation Index (EVI2)	$2.5[(\text{NIR}-\text{R})/(\text{NIR}+2\text{R}+1)]$	(Jiang et al., 2007)
2	Red-edge Normalized Difference Vegetation Index (RENDVI)	$(\text{RE}-\text{R})/(\text{RE}+\text{R})$	(Sims and Gamon, 2002)
3	Normalized Red-Green Difference Index (NGRDVI)	$(\text{G}-\text{R})/(\text{G}+\text{R})$	(Gitelson et al., 2002)
4	NIR-RE Normalized Difference Vegetation index (NIRRENDVI)	$[[\text{(NIR}+\text{RE})/2]-\text{R}]/[[\text{(NIR}+\text{RE})/2]+\text{R}]$	(Xiang et al., 2019)
5	Green Normalized Difference Vegetation Index (GNDVI)	$(\text{NIR}-\text{G})/(\text{NIR}+\text{G})$	(Gitelson and Merzlyak, 1998)
6	Normalized Difference Vegetation Index (NDVI)	$(\text{NIR}-\text{R})/(\text{NIR}+\text{R})$	(Rouse et al., 1974)
7	Red-edge Normalized Difference Vegetation Index (NDVIRE)	$(\text{NIR}-\text{RE})/(\text{NIR}+\text{RE})$	(Sims and Gamon, 2002)
8	Soil Adjusted Vegetation Index (SAVI)	$[(\text{NIR}-\text{R})/(\text{NIR}+\text{R}+\text{L})] (1+\text{L}), \text{L} = 0.5$	(Huete, 1988)

R is the red band, G is the green band, RE is the red-edge band, NIR is the near-infrared band, and L is the soil adjustment factor.

(Ruengvirayudh and Brooks, 2016). Eight VIs were included in the feature selection analysis for plant height estimation. In addition to VIs, UAVH was included for total fresh biomass, fresh stem biomass, fresh leaf biomass, and total plant dry biomass estimation. All analyses were

conducted using the “olsrr” package (Hebbali, 2022) in the R 4.1.3 statistical analysis environment (The R Foundation, Boston, USA).

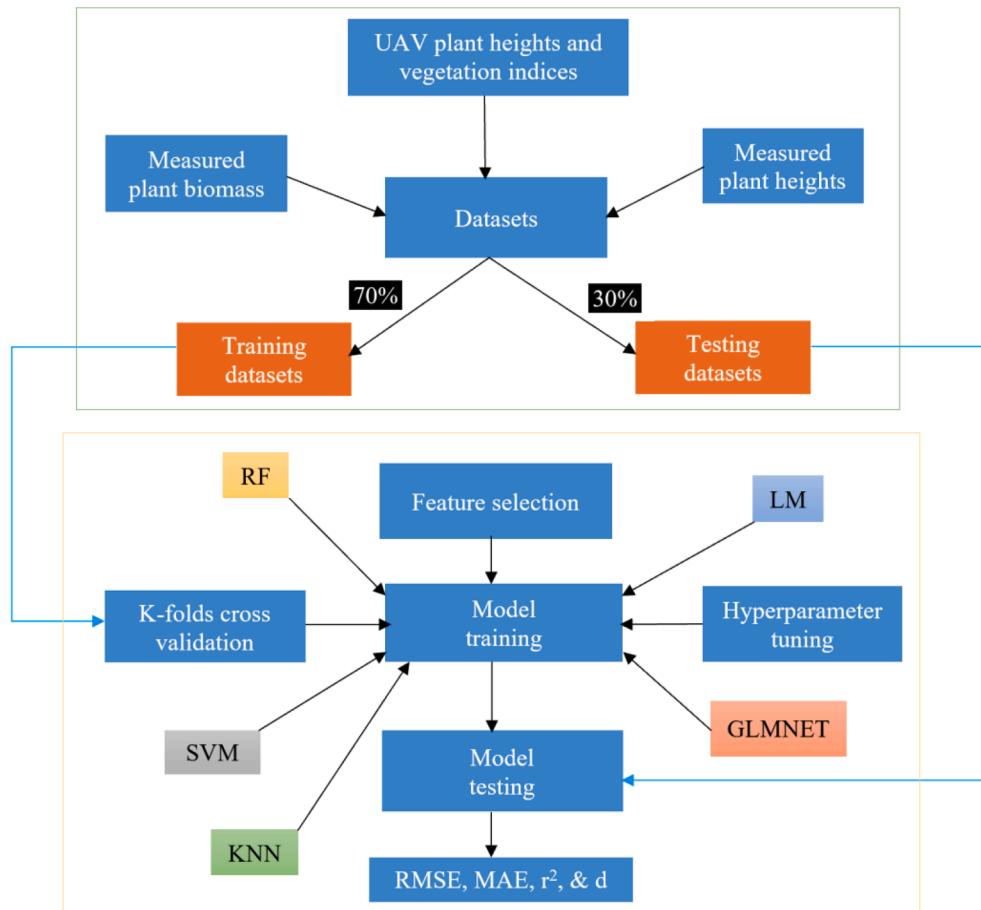
#### 2.7.4. Hyperparameter tuning

This analysis primarily trained LM, GLMNET, RF, kNN, and SVM machine learning algorithms without tuning grids. The optimal default value of each model was then used as an initial point to set the values for the various hyperparameters. The intercept, the hyperparameter for the LM model was held constant for each predicted variable. The random forest model was trained using mtry, which is a hyperparameter that needs tuning. A wide grid of values was used to train the models. Nine mtry values were trained for optimal results. The models’ performance was then investigated to find the optimal mtry value. For the kNN algorithm, the hyperparameter k was used to improve the model outputs. C and sigma hyperparameters were tuned to train the SVM machine learning model. The GLMNET model was tuned using alpha and lambda hyperparameters. The model with the lowest RMSE value was then selected as a prediction model hyperparameter. The “caret” package (Kuhn, 2008) in R 4.1.3 statistical analysis environment was used for hyperparameter tuning.

### 3. Results

#### 3.1. UAV images preprocessing

The georeferencing analysis results showed that georeferencing reduced the average horizontal and vertical RMSE between orthomosaic and independent checkpoints to 0.6 and 1.0 cm. Minimizing



**Fig. 7.** Schematics of data processing and machine learning approaches [RMSE is root mean square error, MAE is mean absolute error, d is an index of agreement, and  $r^2$  is coefficient of determination, LM is linear model, GLMNET is Lasso and elastic-net regularized generalized linear model, RF is the random forest, SVM is support vector machine, and kNN is k-nearest neighbor].

the horizontal and vertical errors between independent checkpoints and ortho-mosaics can improve the reliability of the surface model that is estimated from UAV images later used to estimate plant height (Li et al., 2019). The overall average error between the independent checkpoints and ortho-mosaic images was reduced to 1.2 cm.

### 3.2. Estimating plant height from UAV imagery

The UAV images estimated plant height results showed that UAV images could be effectively used to estimate sweet corn height at different crop growth stages (Fig. 8). Images taken during the early crop growth stage (Dec 8, 2020), resulted in d of 0.66,  $r^2$  of 0.63, and MAE of 1 cm between UAVHs and measured plant height.

The flight that was conducted on Dec 23, 2020, which is a month after planting, yielded better results than the previous flight, with an MAE of 4 cm, d of 0.70, and  $r^2$  of 0.77. The CSM model was able to estimate sweet corn height with relatively high d,  $r^2$ , and RMSE values for specific dates ranging between 0.66 and 0.9, 0.63–0.80, and 1–12 cm, respectively. The average MAE,  $r^2$ , and d between the measured and estimated sweet corn plant heights were 5.4 cm, 0.74, and 0.86, respectively. Overall, the performance of CSM showed improvements as the plant height increased. The result of combined data from all measurement dates showed a strong agreement between measured plant height and UAVHs with the MAE, d, and  $r^2$  of 5 cm, 0.99, and 0.99, respectively (Fig. 8). These findings are in agreement with Ji et al. (2022), Panday et al. (2020), Song and Wang (2019), Chang et al. (2017), and Li et al. (2016), which showed the UAV images estimated plant height had good agreement with the measured height of various crops.

### 3.3. Estimating sweet corn biomass and yield using simple linear regression

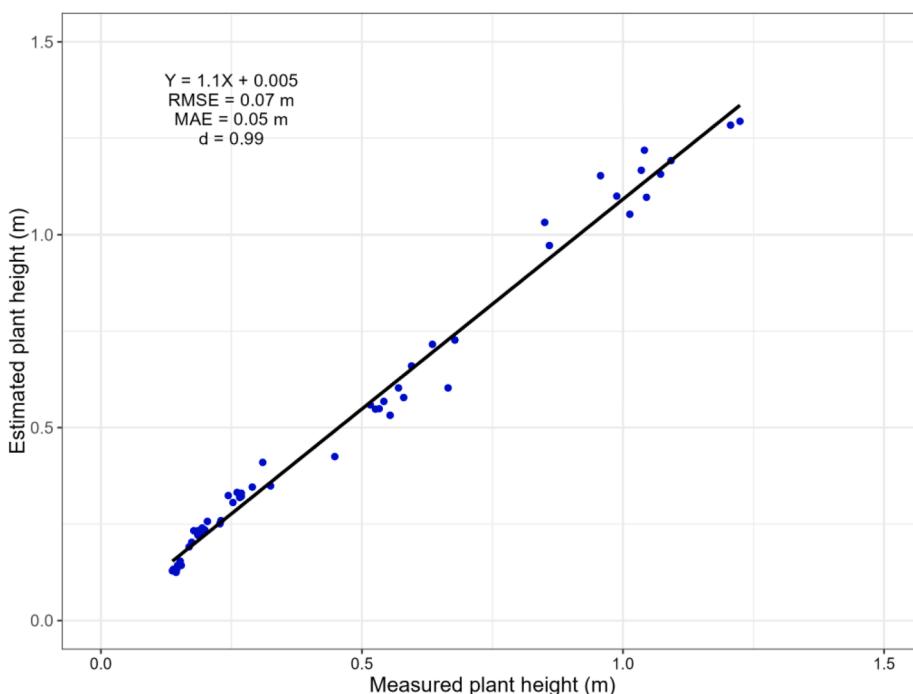
A positive linear relationship between the measured total fresh biomass and UAVH was found with d, adjusted  $r^2$  and MAE of 0.91, 0.88, and 0.24 kg m<sup>-2</sup>, respectively. The adjusted  $r^2$ , d, and MAE of 0.90, 0.65, and 0.33 kg m<sup>-2</sup> were found between measured fresh leaf biomass

and UAVH. Comparable results were observed for fresh stem biomass and total dry biomass, where the d, adjusted  $r^2$ , and MAE were 0.94, 0.87, and 0.20 kg m<sup>-2</sup> and 0.66, 0.78, and 0.36 kg m<sup>-2</sup>, respectively. Similar findings were reported by several researchers. Bendig et al. (2014) revealed that the correlation between barley (*Hordeum vulgare L.*) height from CSMs and fresh biomass was strong ( $r^2 = 0.81$ ). Kim et al. (2018) found comparable results, where the UAV images estimated Chinese cabbage (*Brassica rapa*) and white radish (*Raphanus sativus L.*) heights were highly correlated with corresponding fresh biomass. The cross-validation results indicated that the UAVH could accurately predict the sweet corn fresh total biomass with d and MAE of 0.75 and 0.24 kg m<sup>-2</sup> and dry total biomass with d and MAE of 0.7 and 0.36 kg m<sup>-2</sup>. A study conducted by Lussem et al. (2019a) showed that the use of UAV-based height for estimating temperate grassland biomass provided good results for both fresh and dry biomass. Brocks and Bareth (2018) noted the same, where the crop surface model's plot-wise plant heights could be used to estimate both fresh and dry biomass of barley. Overall, the result revealed that the correlations between the UAV images estimated and measured sweet corn fresh and dry biomass were good (Fig. 9).

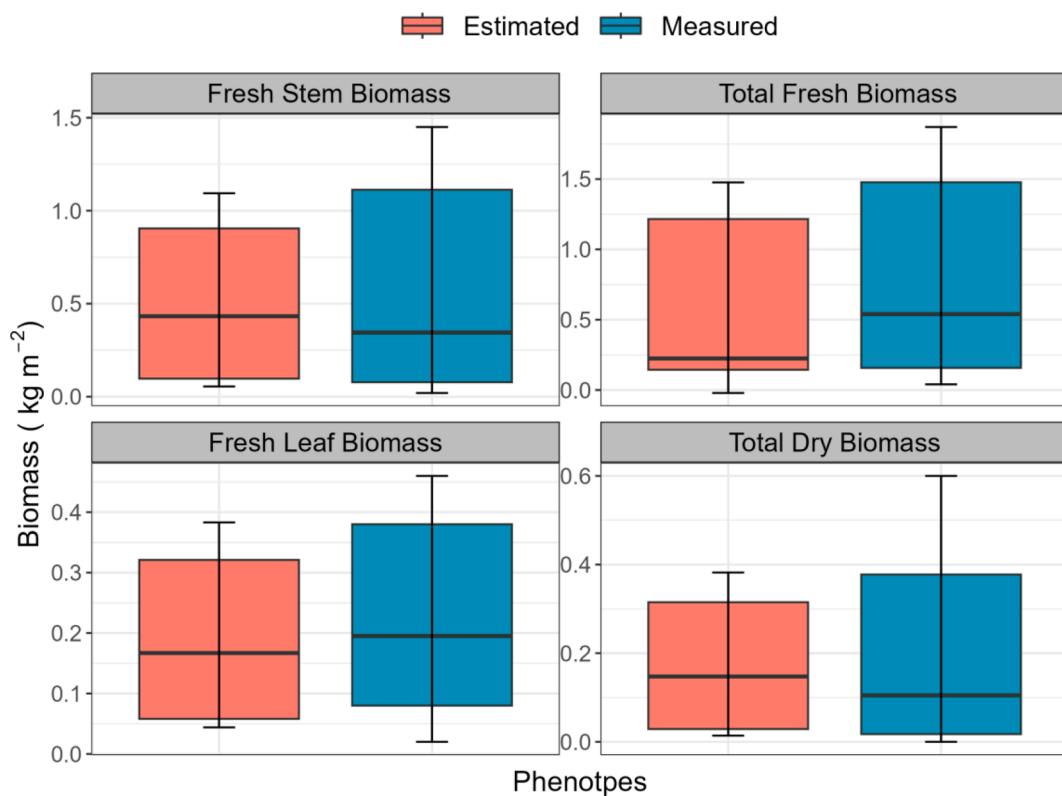
A positive correlation between the measured yield and UAVH was found with d, adjusted  $r^2$ , and MAE of 0.84, 0.66, and 0.07 kg m<sup>-2</sup>, respectively. Earlier studies also have shown that UAVH can be used for estimating crop yield. A study by Panday et al. (2020) reported that the wheat height derived from CSM positively correlated with grain weight. Feng et al. (2019) findings also indicated that the UAV-based remote sensing system that was equipped with a digital camera was capable of monitoring growth and estimating cotton yield.

### 3.4. Vegetation indices (VIs)

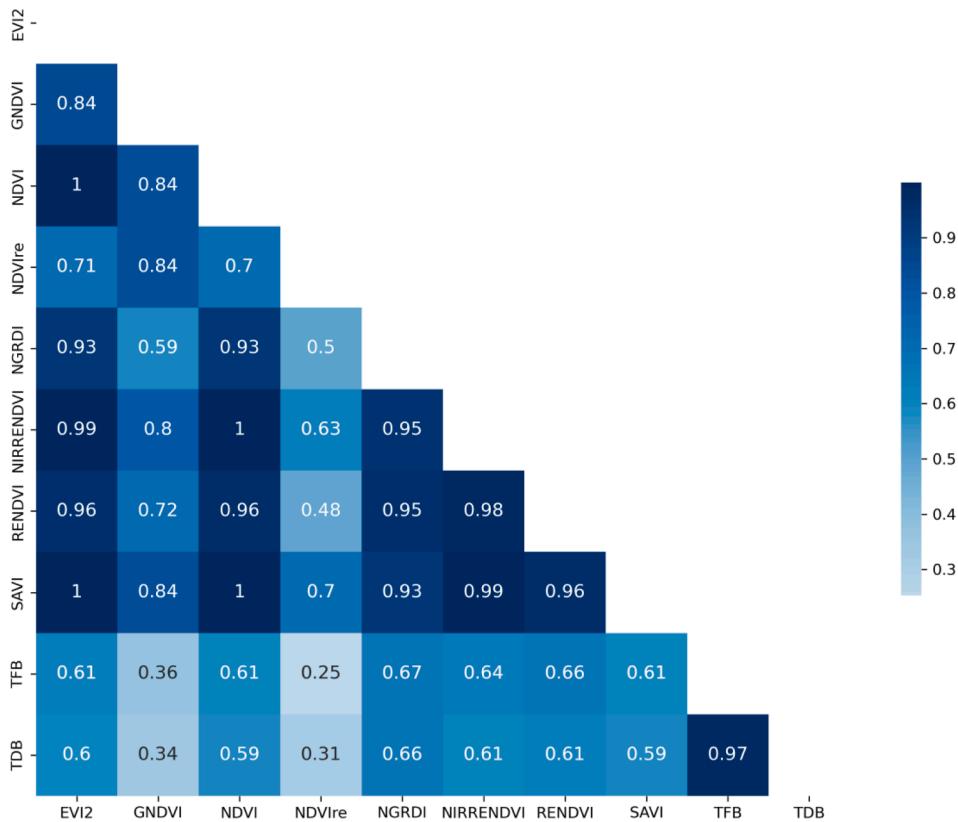
The results correlating measured plant height with the VIs were mixed. A good correlation was observed between plant height and NIRRENDVI, RENDVI, and NGRDI indices (Fig. 10). The lowest correlation was found with NDVI<sub>re</sub>, having a correlation coefficient of 0.29. However, the UAV images estimated plant height (UAVH) correlation with these three VIs was less than that of the measured plant height. Thenkabail et al. (1994) reported that the relationship between mid-



**Fig. 8.** Scatter plot between measured plant heights and UAV images estimated plant heights [RMSE is root mean square error, MAE is mean absolute error, and d is an index of agreement].



**Fig. 9.** Box plots for estimated total fresh biomass, total dry biomass, fresh stem biomass, and fresh leaf biomass against measured respective biomass.



**Fig. 10.** Correlation between vegetation indices and biomass (fresh and dry) [EVI2 is Enhanced Vegetation Index, GNDVI is Green Normalized Difference Vegetation Index, NDVI is Normalized Difference Vegetation Index, NDVIRE is Red-edge Normalized Difference Vegetation Index, NGRDI is Normalized Red-Green Difference Index, NIRRENDVI is NIR-RE normalized difference vegetation index, RENDVI is Red-edge Normalized Difference Vegetation Index, and SAVI is Soil Adjusted Vegetation Index, TFB is total fresh biomass, and TDB is total dry biomass]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

infrared-based vegetation indices and plant height of soybean (*Glycine max*) and corn was generally positive. Payero et al. (2004) found that the relationship between plant height and VIs of alfalfa (*Medicago sativa*) and grass (*Poaceae*) was good. They also noted that the changes in plant height during the start of the growing cycle were very sensitive to the changes in VIs.

The biomass had a good correlation with the six VIs, namely the NDVI, RENDVI, NIRRENDVI, NGRDI, RENDVI, and SAVI (Fig. 10). Biomass correlations with the remaining indices, namely NDVIRE and GNDVI, were weak. The NGRDI had the highest correlations with biomass (total fresh biomass, fresh stem biomass, fresh leaf biomass, and total dry biomass) than other indices with correlation coefficients of 0.67, 0.66, 0.70, and 0.66, respectively. The NDVIRE had the lowest correlation with biomass with correlation coefficients of 0.25, 0.24, 0.29, and 0.31, respectively (Fig. 10). The fresh leaf biomass had a strong correlation with RENDVI. Although vegetation indices based on spectral features are promising for monitoring biomass, their accuracy and suitability vary from index to index (Lussem et al., 2019b). Niu et al. (2019) reported that UAV images-based vegetation indices can be used to estimate the biomass of maize at the above-ground level; the NGRDI had positive correlations with both dry and fresh above ground biomass (AGB). Zheng et al. (2022) found that red-edge-related VIs were the most influential VIs when it came to predicting biomass. In addition, the correlation result among VIs revealed that most indices had strong correlations with each other ( $r > 0.70$ ) (Fig. 10). The presence of strong correlations between predictors could limit the number of features that are used in ML models. Whitmire et al. (2021) found that the correlation-based feature selection was the most accurate feature selection method for predicting the biomass and yield of alfalfa. Therefore, in this study, the correlation between the vegetation indices and the correlation of vegetation indices with biomass was utilized in pre-screening predictors in machine learning plant height and biomass prediction.

### 3.5. Feature selection

The results of a feature selection analysis revealed that the model that included EVI2, NGRDI, NIRRENDVI, GNDVI, and SAVI vegetation indices was the best model for plant height estimation with an adjusted  $r^2$  and MSEP values of 0.64 and 2.59. The model that only consists of UAVH was regarded as the most accurate model for estimating total fresh plant biomass and fresh leaf biomass with adjusted  $r^2$  and MSEP values of 0.88 and 1.94 and 0.90 and 0.095 (Table 2), respectively. Usually, fewer data-demanding models are preferred over feature-intensive ones, which can increase the demand for computational resources and cause issues with the dimensionality (Cai et al., 2018). The

best estimator of the fresh stem biomass and total dry biomass of sweet corn plants was the model that included EVI2, GNDVI, NGRDI, NIRRENDVI, NDVIRE, and UAVH features (Table 2). Maimaitijiang et al. (2020) showed that adding information about the canopy structure to spectral features can reduce the background effect of the soil and can improve the performance of ML models. Using plant height parameters along with vegetation indices has improved the performance of the models for fresh stem and total dry biomass estimations. Several researchers have previously reported the same result. Zheng et al. (2022) stated that the combination of the VIs and canopy geometric parameters obtained from UAV imagery was able to provide an effective estimate of dry biomass for strawberry (*Fragaria × ananassa*) fields. Viljanen et al. (2018) showed that the best estimate of biomass was made using the combination of the VIs and crop height. Tilly et al. (2015) revealed the potential of remotely-sensed plant parameters such as plant height and VIs for the prediction of barley (*Hordeum vulgare L*) fresh and dry biomass.

### 3.6. Hyperparameter tuning

The final mtry values used to estimate the plant height, fresh leaf biomass, total fresh biomass, fresh stem biomass, and total dry biomass were 7, 7, 2, 3, and 7, respectively. Table 3 shows the hyperparameters grids and best hyperparameters values used for plant height and biomass estimation. The final k values used to estimate these parameters using the kNN model were 6, 6, 6, 4, and 4, respectively (Table 3). For the SVM, the grid value used for sigma hyperparameter optimization ranged between 0.1 and 0.5, while the C hyperparameter grid value ranged between 0.05 and 5. The final sigma and C hyperparameter values used to estimate aforesaid sweet corn phenotypic parameters using SVM were 0.3, 0.3, 0.5, 0.5, and 0.4 and 3.55, 4.55, 3.05, 4.55, and 2.05, respectively (Table 3). Several researchers noted that optimizing the hyperparameters of machine learning algorithms can help improve their performance (Paudel et al., 2021; Pranga et al., 2021; Probst et al., 2019; Revenga et al., 2022; Schratz et al., 2019).

### 3.7. Training and test results

#### 3.7.1. Biomass estimation using machine learning models

The kNN, SVM, and RF were able to accurately estimate the total fresh biomass of sweet corn (Table 4). The testing  $r^2$  results of the kNN, SVM, and RF were 0.86, 0.85, and 0.78, respectively. The same models performed well for fresh leaf biomass estimation. The kNN model was able to estimate the fresh leaf biomass with an accuracy of 0.92 ( $r^2$ ) and 70 g m<sup>-2</sup> (RMSE).

The SVM model estimated the fresh leaf biomass with an accuracy of

**Table 2**

Best input features for plant height and biomass estimation based on the best subsets regression feature selection method.

Model	Fresh leaf biomass	Total fresh biomass	Fresh stem biomass	Total dry biomass	Plant height
1	UAVH <sup>a</sup>	UAVH <sup>a</sup>	UAVH	UAVH	RENDVI
2	EVI2, UAVH	GNDVI, UAVH	GNDVI, UAVH	NDVIRE, UAVH	NIRRENDVI, NDVI
3	NDVI, SAVI, UAVH	GNDVI, NDVIRE, UAVH	GNDVI, NDVIRE, UAVH	GNDVI, NDVIRE, UAVH	EVI2, NIRRENDVI, SAVI
4	EVI2, GNDVI, NGRDI, UAVH	EVI2, GNDVI, NGRDI, UAVH	GNDVI, NDVI, NIRRENDVI, UAVH	GNDVI, NDVI, RENDVI, UAVH	EVI2, NIRRENDVI, NDVI, SAVI
5	EVI2, GNDVI, NGRDI, SAVI, UAVH	EVI2, GNDVI, NGRDI, SAVI, UAVH	GNDVI, NDVIRE, NGRDI, RENDVI, UAVH	GNDVI, NDVIRE, NGRDI, RENDVI, UAVH	EVI2, NGRDI, NIRRENDVI, GNDVI, SAVI <sup>a</sup>
6	EVI2, GNDVI, NDVI NDVIRE, NGRDI, UAVH	EVI2, GNDVI, NDVI NDVIRE, NGRDI, NIRRENDVI, UAVH	EVI2, GNDVI, NDVI NDVIRE, NGRDI, NIRRENDVI, UAVH <sup>a</sup>	EVI2, GNDVI, NDVI NDVIRE, NGRDI, NIRRENDVI, UAVH <sup>a</sup>	EVI2, NGRDI, NIRRENDVI, GNDVI, NDVI, SAVI
7	EVI2, GNDVI, NDVI NDVIRE, NGRDI, RENDVI UAVH	EVI2, GNDVI, NDVI NDVIRE, NGRDI, RENDVI, UAVH	EVI2, GNDVI, NDVI NDVIRE, NGRDI, RENDVI, UAVH	EVI2, GNDVI, NDVI NDVIRE, NGRDI, RENDVI, UAVH	EVI2, NGRDI, NIRRENDVI, GNDVI, NDVI, NDVIRE, SAVI
8	EVI2, GNDVI, NDVI, NDVIRE, NGRDI NIRRENDVI, RENDVI, UAVH	EVI2, GNDVI, NDVI, NDVIRE, NGRDI, NIRRENDVI, RENDVI, UAVH	EVI2, GNDVI, NDVI, NDVIRE, NGRDI, NIRRENDVI, RENDVI, SAVI, UAVH	EVI2, GNDVI, NDVI, NDVIRE, NGRDI, NIRRENDVI, RENDVI, SAVI, UAVH	EVI2, RENDVI, NGRDI, NIRRENDVI, GNDVI, NDVI, NDVIRE, SAVI
9	EVI2, GNDVI, NDVI, NDVIRE, NGRDI, NIRRENDVI, RENDVI, SAVI, UAVH	EVI2, GNDVI, NDVI, NDVIRE, NGRDI, NIRRENDVI, RENDVI, SAVI, UAVH	EVI2, GNDVI, NDVI, NDVIRE, NGRDI, NIRRENDVI, RENDVI, SAVI, UAVH	EVI2, GNDVI, NDVI, NDVIRE, NGRDI, NIRRENDVI, RENDVI, SAVI, UAVH	EVI2, GNDVI, NDVI, NDVIRE, NGRDI, NIRRENDVI, RENDVI, SAVI

<sup>a</sup> selected model for plant height and biomass estimation.

**Table 3**  
Tuning grids and optimal values.

Model	Hyperparameters	Tuning grids	Best hyperparameters				
			Plant height	Fresh leaf biomass	Total fresh biomass	Fresh stem biomass	Total dry biomass
LM	intercept	constant	constant	constant	constant	constant	constant
	alpha	0.1–1	0.6	0.1	0.4	0.1	0.6
	lambda	0.001–0.01	0.001	0.001	0.002	0.001	0.001
RF	mtry	1–9	7	7	2	3	7
	K	2–10	6	6	6	4	4
kNN	sigma	0.1–0.5	0.3	0.3	0.5	0.5	0.4
	C	0.05–5	3.55	4.55	3.05	4.55	2.05

**Table 4**  
Training and testing results of biomass and plant height.

Parameters		Evaluation	LM	GLMNET	RF	kNN	SVM
Fresh leaf biomass	Training	$r^2$	0.91	0.93	0.9	0.93	0.92
		RMSE	0.06	0.06	0.05	0.04	0.04
		MAE	0.06	0.05	0.04	0.03	0.03
	Testing	$r^2$	0.83	0.8	0.87	0.92	0.92
		RMSE	0.07	0.07	0.06	0.07	0.06
		MAE	0.06	0.06	0.03	0.04	0.04
Total fresh biomass	Training	$r^2$	0.94	0.93	0.95	0.95	0.96
		RMSE	0.32	0.26	0.2	0.19	0.18
		MAE	0.29	0.14	0.16	0.14	0.14
	Testing	$r^2$	0.76	0.71	0.78	0.86	0.85
		RMSE	0.31	0.35	0.32	0.26	0.25
		MAE	0.29	0.25	0.29	0.12	0.25
Fresh stem biomass	Training	$r^2$	0.93	0.85	0.93	0.96	0.85
		RMSE	0.26	0.24	0.16	0.15	0.21
		MAE	0.24	0.21	0.12	0.11	0.17
	Testing	$r^2$	0.77	0.76	0.77	0.81	0.73
		RMSE	0.24	0.25	0.25	0.23	0.28
		MAE	0.22	0.18	0.13	0.12	0.18
Total dry biomass	Training	$r^2$	0.93	0.92	0.93	0.95	0.91
		RMSE	0.1	0.11	0.08	0.09	0.09
		MAE	0.09	0.1	0.07	0.06	0.08
	Testing	$r^2$	0.69	0.68	0.76	0.84	0.85
		RMSE	0.09	0.1	0.09	0.07	0.07
		MAE	0.08	0.08	0.04	0.03	0.04
Plant Height	Training	$r^2$	0.89	0.88	0.90	0.96	0.95
		RMSE	0.31	0.29	0.29	0.29	0.26
		MAE	0.26	0.28	0.23	0.23	0.20
	Testing	$r^2$	0.79	0.81	0.73	0.54	0.65
		RMSE	0.22	0.22	0.3	0.3	0.26
		MAE	0.17	0.22	0.20	0.22	0.19
		d	0.90	0.79	0.78	0.81	0.87

LM is the linear model, GLMNET is Lasso and elastic-net regularized generalized linear models, RF is the random forest, kNN is k nearest neighbor, SVM is the support vector machine, RMSE is the root mean square error, and  $r^2$  is the coefficient of determination.

0.92  $r^2$  and 60 g m<sup>-2</sup> RMSE, while the RF model estimated the fresh leaf biomass with 0.87  $r^2$  and 60 g m<sup>-2</sup> RMSE. The  $r^2$  revealed that they were able to provide an accurate estimate of the total fresh biomass with  $r^2$  of 0.78, 0.85, and 0.85, respectively. Likewise, all the models performed well in estimating fresh stem biomass. However, the training and testing results of GLMNET and LM models did not reflect the actual performance (Table 4). The model's performance decreased when it came to estimating unseen data of the target variable. The maximum estimation performance was observed with the kNN model with training and testing  $r^2$  of 0.84 and 0.81, respectively. Similar scenarios were observed for total dry biomass estimation, where the kNN, RF, and SVM models performed better than the GLMNET and LM models. The testing  $r^2$  results of the kNN, RF, and SVM were 0.76, 0.84, and 0.85, respectively. This result agrees with several research reports, where machine learning models can estimate fresh and dry biomass using UAV images-based vegetation indices and crop height information. Liu et al. (2022) identified Gaussian process regression (GPR), SVM, and RF were the best ML

models to estimate the above-ground biomass of potatoes (*Solanum tuberosum*) using UAV hyperspectral images extracted canopy spectral information and crop height. Osco et al. (2020) and Han et al. (2019) showed that the RF model can accurately estimate the agronomic variables of maize plants. Yue et al. (2018) also revealed RF is very robust against noise. It is ideal for dealing with repeated observations that involve data from sensors. Whitmire et al. (2021) indicated RF and kNN machine-learning models can confidently estimate alfalfa biomass and crop yields even on small datasets with a few features. Although the  $r^2$  results of the LM and GLMNET models indicated that the models performed well in the training, they did not generalize the unseen datasets well enough, where the performances of the models during testing were less than those training (Table 4). Overall, the SVM and kNN models performed well in estimating the plant's fresh and dry biomass. However, the SVM performed better than the kNN when it came to estimating most of the biomass.

### 3.7.2. Plant height estimation using machine learning models

The GLMNET and LM models performed well in estimating plant height. The GLMNET model had an estimation accuracy of 0.81 and an RMSE of 2.2 cm, while the LM model estimated plant height with an  $r^2$  of 0.79 and an RMSE of 2.2 cm. Nickmilder et al. (2021) reported that the GLMNET model was the best-performing model for compressed sward height estimation of Walloon pastures. The RF model performed better for plant height estimation compared to the SVM and kNN models. This is in agreement with Teodoro's (2021) findings, where the RF model presented a strong estimation capacity for the plant height of Soybean. Overall, the GLMNET machine learning model outperformed all other models when it comes to estimating sweet corn plant heights. Besides, compared to simple linear regression models, machine learning models performed better when it came to predicting sweet corn plant biomass and plant height.

The results of this research paper showed that the integration of UAV imaging and machine learning techniques can significantly improve plant phenotyping accuracy and efficiency. By utilizing high-resolution aerial images and machine learning algorithms, we were able to accurately measure plant height, biomass, and other important phenotypic traits. These findings have important implications for the future of plant breeding and crop management, as they provide a cost-effective and non-invasive way to rapidly assess plant growth and development. High-throughput phenotyping based on high-resolution aerial imaging coupled with machine learning provides economic benefits in terms of enhancing efficiency, cost savings, improved accuracy and precision, early issue detection, and scalability (Guerrero-Ibañez et al., 2021; Heidari et al., 2023; Rehman et al., 2022; Tian et al., 2020). These economic benefits make these techniques the primary alternative for scientists, growers, and crop managers to improve productivity and optimize resource allocation in the agricultural sector (Amarasingam et al., 2022; Boursianis et al., 2022; Chawade et al., 2019; Olson and Anderson, 2021; Xie and Yang, 2020). Overall, this study highlights the exciting potential of UAV imaging and machine learning in the field of plant phenotyping.

## 4. Conclusion

When it comes to monitoring plant phenotypes, it is important to take immediate action to address any issues before they lead to yield loss. Unfortunately, field sampling is often very time-consuming and costly. Also, some techniques are destructive. Non-destructive methods are starting to emerge as viable tools for monitoring crop phenological traits. Applications of UAV-based imaging for plant phenotyping and machine learning are two fields growing rapidly. Findings from this study demonstrated that UAV-based multispectral imaging and machine learning algorithms can be effectively used to estimate sweet corn height, biomass, and yield with reasonable accuracy. Using UAV imaging and machine learning techniques, statistical performance measures, such as the  $d$ ,  $r^2$ , and RMSE, showed a strong correlation between measured and estimated plant parameters. Specifically, the kNN and SVM models outperformed the other machine learning models when it comes to estimating plant biomass, while the GLMNET performed well in estimating plant height. Overall, the results suggest that UAV imaging and machine learning techniques can be used to monitor and estimate plant yield and yield components, which in turn could be used to make informed decisions at plot and field levels. Further studies are also needed to refine these techniques' accuracy and explore their potential for real-world applications. The applicability of the UAV imaging and machine learning techniques for monitoring plant phenotypes under different environmental and management conditions should also be evaluated.

## 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.

## Data availability

Data will be made available on request.

## Acknowledgment

Funding for this research was provided by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under award number 2020-67021-31965, and the Hatch Project No. FLA-TRC-005751. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture. The authors also thank Dr. Fikadu Welidehanna and Mr. Christian Bartell for their invaluable assistance in gathering data throughout the experiment.

## References

- Adelabu, S., Mutanga, O., Adam, E., 2014. Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels. *ISPRS J. Photogramm. Remote Sens.* 95, 34–41. <https://doi.org/10.1016/j.isprsjprs.2014.05.013>.
- Ahamed, T., Tian, L., Zhang, Y., Ting, K.C., 2011. A review of remote sensing methods for biomass feedstock production. *Biomass Bioenergy* 35, 2455–2469. <https://doi.org/10.1016/j.biombioe.2011.02.028>.
- Alamoodi, A.H., Zaidan, B.B., Zaidan, A.A., Albahri, O.S., Chen, J., Chyad, M.A., Garfan, S., Aleesa, A.M., 2021. Machine learning-based imputation soft computing approach for large missing scale and non-reference data imputation. *Chaos Solitons Fractals* 151, 111236. <https://doi.org/10.1016/j.chaos.2021.111236>.
- Ali, E.M., Ahmed, M.M., Wulff, S.S., 2019. Detection of critical safety events on freeways in clear and rainy weather using SHRP2 naturalistic driving data: Parametric and non-parametric techniques. *Saf. Sci.* 119, 141–149. <https://doi.org/10.1016/j.ssci.2019.01.007>.
- Alonzo, L.M.B., Chioson, F.B., Co, H.S., Bugtai, N.T., Baldovino, R.G., 2018. A Machine Learning Approach for Coconut Sugar Quality Assessment and Prediction, in: 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM). Presented at the 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM), pp. 1–4. <https://doi.org/10.1109/HNICEM.2018.8666315>.
- Ampatzidis, Y., Partel, V., 2019. UAV-Based High Throughput Phenotyping in Citrus Utilizing Multispectral Imaging and Artificial Intelligence. *Remote Sens. (Basel)* 11, 410. <https://doi.org/10.3390/rs11040410>.
- Ampatzidis, Y., Partel, V., Costa, L., 2020. Agroview: Cloud-based application to process, analyze and visualize UAV-collected data for precision agriculture applications utilizing artificial intelligence. *Comput. Electron. Agric.* 174, 105457 <https://doi.org/10.1016/j.compag.2020.105457>.
- Anthony, D., Elbaum, S., Lorenz, A., Detweiler, C., 2014. On crop height estimation with UAVs. In: In: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. Presented at the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014). IEEE, Chicago, IL, USA, pp. 4805–4812. <https://doi.org/10.1109/IROS.2014.6943245>.
- Atefi, A., Ge, Y., Pitla, S., Schnable, J., 2021. Robotic technologies for high-throughput plant phenotyping: Contemporary reviews and future perspectives. *Front. Plant Sci.* 12.
- Barrientos, A., Colorado, J., del Cerro, J., Martinez, A., Rossi, C., Sanz, D., Valente, J., 2011. Aerial remote sensing in agriculture: A practical approach to area coverage and path planning for fleets of mini aerial robots. *J. Field Rob.* 28, 667–689. <https://doi.org/10.1002/rob.20403>.
- Battude, M., Al Bitar, A., Morin, D., Cros, J., Huc, M., Marais Sicre, C., Le Dantec, V., Demarez, V., 2016. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sens. Environ.* 184, 668–681. <https://doi.org/10.1016/j.rse.2016.07.030>.
- Becker-Reshef, I., Justice, C., Barker, B., Humber, M., Rembold, F., Bonifacio, R., Zappacosta, M., Budde, M., Magadire, T., Shitote, C., Pound, J., Constantino, A., Nakalembe, C., Mwangi, K., Sobue, S., Newby, T., Whitcraft, A., Jarvis, I., Verdin, J., 2020. Strengthening agricultural decisions in countries at risk of food insecurity: The GEOGLAM Crop Monitor for Early Warning. *Remote Sens. Environ.* 237, 111553 <https://doi.org/10.1016/j.rse.2019.111553>.
- Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., Bareth, G., 2014. Estimating biomass of barley using Crop Surface Models (CSMs) derived from UAV-based RGB imaging. *Remote Sens. (Basel)* 6, 10395–10412. <https://doi.org/10.3390/rs61110395>.
- Berrar, D., 2018. Cross-Validation. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>.
- Boulesteix, A.-L., Janitz, S., Kruppa, J., König, I.R., 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *WIREs Data Min. Knowl. Discovery* 2, 493–507. <https://doi.org/10.1002/widm.1072>.

- Brocks, S., Bareth, G., 2018. Estimating barley biomass with crop surface models from oblique RGB imagery. *Remote Sens. (Basel)* 10, 268. <https://doi.org/10.3390/rs10020268>.
- Cai, J., Luo, J., Wang, S., Yang, S., 2018. Feature selection in machine learning: A new perspective. *Neurocomputing* 300, 70–79. <https://doi.org/10.1016/j.neucom.2017.11.077>.
- Candiego, S., Remondino, F., De Giglio, M., Dubbini, M., Gattelli, M., 2015. Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sens. (Basel)* 7, 4026–4047. <https://doi.org/10.3390/rs70404026>.
- Chang, A., Jung, J., Maeda, M.M., Landivar, J., 2017. Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). *Comput. Electron. Agric.* 141, 232–237. <https://doi.org/10.1016/j.compag.2017.07.008>.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20, 273–297. <https://doi.org/10.1007/BF00994018>.
- Cosenza, D.N., Packalen, P., Maltamo, M., Varvia, P., Räty, J., Soares, P., Tomé, M., Strunk, J.L., Korhonen, L., 2022. Effects of numbers of observations and predictors for various model types on the performance of forest inventory with airborne laser scanning. *Can. J. For. Res.* 52, 385–395. <https://doi.org/10.1139/cjfr-2021-0192>.
- Crampton, J.W., 2016. Assemblage of the vertical: commercial drones and algorithmic life. *Geographica Helvetica* 71, 137–146. <https://doi.org/10.5194/gb-71-137-2016>.
- Cwiakala, P., Kocierz, R., Punia, E., Nędzka, M., Mamczarz, K., Niewiem, W., Wiącek, P., 2018. Assessment of the possibility of using Unmanned Aerial Vehicles (UAVs) for the documentation of hiking trails in alpine areas. *Sensors* 18, 81. <https://doi.org/10.3390/s18010081>.
- de Vlaming, R., Groenen, P.J.F., 2015. The current and future use of ridge regression for prediction in quantitative genetics. *Biomed Res. Int.* 2015, e143712.
- Deery, D.M., Jones, H.G., 2021. Field Phenomics: Will It Enable Crop Improvement? *Plant Phenomics* 2021, 2021/9871989. <https://doi.org/10.34133/2021/9871989>. Deery.
- Delavarpour, N., Koparan, C., Nowatzki, J., Bajwa, S., Sun, X., 2021. A technical study on UAV characteristics for precision agriculture applications and associated practical challenges. *Remote Sens. (Basel)* 13, 1204. <https://doi.org/10.3390/rs13061204>.
- Delen, D., Oztekin, A., Kong, Z. (James), 2010. A machine learning-based approach to prognostic analysis of thoracic transplants. *Artificial Intelligence in Medicine* 49, 33–42. <https://doi.org/10.1016/j.artmed.2010.01.002>.
- Dicke, D., Jacobi, J., Büchse, A., 2012. Quantifying herbicide injuries in maize by use of remote sensing. *Julius-Kühn-Archiv* 1, 199–205.
- Dittmar, D.P., Bryant, T., Crawford, H., 2021. Handbook Design and Composition 579.
- Doughty, C.L., Cavanaugh, K.C., 2019. Mapping coastal wetland biomass from high resolution Unmanned Aerial Vehicle (UAV) imagery. *Remote Sens. (Basel)* 11, 540. <https://doi.org/10.3390/rs11050540>.
- Easterday, K., Kislik, C., Dawson, T.E., Hogan, S., Kelly, M., 2019. Remotely sensed water limitation in vegetation: Insights from an experiment with Unmanned Aerial Vehicles (UAVs). *Remote Sens. (Basel)* 11, 1853. <https://doi.org/10.3390/rs11161853>.
- Esquerdo, J.C.D.M., Zullo Júnior, J., Antunes, J.F.G., 2011. Use of NDVI/AVHRR time-series profiles for soybean crop monitoring in Brazil. *Int. J. Remote Sens.* 32, 3711–3727. <https://doi.org/10.1080/01431161003764112>.
- Feng, A., Zhang, M., Sudduth, K.A., Vories, E.D., Zhou, J., 2019. Cotton yield estimation from UAV-based plant height. *Trans. ASABE* 62, 393–404. <https://doi.org/10.13031/trans.13067>.
- Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22.
- 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. [https://doi.org/10.1016/S0034-4257\(01\)00289-9](https://doi.org/10.1016/S0034-4257(01)00289-9).
- Gitelson, A., Merzlyak, M., 1998. Remote sensing of chlorophyll concentration in higher plant leaves. *Adv. Space Res. - ADV. SPACE RES.* 22, 689–692. [https://doi.org/10.1016/S0273-1177\(97\)01133-2](https://doi.org/10.1016/S0273-1177(97)01133-2).
- Goel, P.K., Prasher, S.O., Landry, J.A., Patel, R.M., Bonnell, R.B., Viau, A.A., Miller, J.R., 2003. Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn. *Comput. Electron. Agric.* 38, 99–124. [https://doi.org/10.1016/S0168-1699\(02\)00138-2](https://doi.org/10.1016/S0168-1699(02)00138-2).
- Guo, Y., Jia, X., Paull, D., Zhang, J., Farooq, A., Chen, X., Islam, Md.N., 2019. A Drone-Based Sensing System to Support Satellite Image Analysis for Rice Farm Mapping, in: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, pp. 9376–9379. <https://doi.org/10.1109/IGARSS.2019.8898638>.
- Guo, Y., Wang, H., Wu, Z., Wang, S., Sun, H., Senthilnath, J., Wang, J., Robin Bryant, C., Fu, Y., 2020. Modified red blue vegetation index for chlorophyll estimation and yield prediction of maize from visible images captured by UAV. *Sensors* 20, 5055. <https://doi.org/10.3390/s20185055>.
- Gupta, A., Afrin, T., Scully, E., Yodo, N., 2021. Advances of UAVs toward future transportation: The state-of-the-art, challenges, and opportunities. *Future Transportation* 1, 326–350. <https://doi.org/10.3390/futuretransp1020019>.
- Han, X., Thomasson, J., Bagnall, G., Pugh, N., Horne, D., Rooney, W., Jung, J., Chang, A., Malambo, L., Popescu, S., Gates, I., Cope, D., 2018. Measurement and calibration of plant-height from fixed-wing UAV images. *Sensors* 18, 4092. <https://doi.org/10.3390/s18124092>.
- Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., Li, Z., Yang, X., 2019. Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods* 15, 10. <https://doi.org/10.1186/s13007-019-0394-z>.
- Hebbali, A., 2022. olsrr. Rsquared Academy (<https://olsrr.rsquaredacademy.com/license-text>).
- Holman, F.H., Riche, A.B., Michalski, A., Castle, M., Wooster, M.J., Hawkesford, M.J., 2016. High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. *Remote Sens. (Basel)* 8, 1031. <https://doi.org/10.3390/rs8121031>.
- Huang, Y., Thomson, S.J., Hoffmann, W.C., Lan, Y., Fritz, B.K., 2013. Development and prospect of unmanned aerial vehicle technologies for agricultural production management. *Int. J. Agric. Biol. Eng.* 6, 1–10. <https://doi.org/10.2516/ijabe.v6i3.900>.
- Huang, Y., Chen, Z., Yu, T., Huang, X., Gu, X., 2018a. Agricultural remote sensing big data: Management and applications. *J. Integr. Agric.* 17, 1915–1931. [https://doi.org/10.1016/S2095-3119\(17\)61859-8](https://doi.org/10.1016/S2095-3119(17)61859-8).
- Huang, Y., Reddy, K.N., Fletcher, R.S., Pennington, D., 2018b. UAV low-altitude remote sensing for precision weed management. *Weed Technol.* 32, 2–6. <https://doi.org/10.1017/wet.2017.89>.
- Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25, 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X).
- Ismail, R., Mutanga, O., 2010. A comparison of regression tree ensembles: Predicting Sirex nootilio induced water stress in Pinus patula forests of KwaZulu-Natal, South Africa. *Int. J. Appl. Earth Obs. Geoinf.* 12, S45–S51. <https://doi.org/10.1016/j.jag.2009.09.004>.
- Ji, Y., Chen, Z., Cheng, Q., Liu, R., Li, M., Yan, X., Li, G., Wang, D., Fu, L., Ma, Y., Jin, X., Zong, X., Yang, T., 2022. Estimation of plant height and yield based on UAV imagery in faba bean (*Vicia faba* L.). *Plant Methods* 18, 26. <https://doi.org/10.1186/s13007-022-00861-7>.
- Jiang, Z., Huete, A., Kim, Y., Didan, K., 2007. 2-Band enhanced vegetation index without a blue band and its application to AVHRR data. *Proc. SPIE-Int. Soc. Opt. Eng.* 6679 <https://doi.org/10.1117/12.734933>.
- Jiménez, S., Angeles-Valdez, D., Villicana, V., Reyes-Zamorano, E., Alcalá-Lozano, R., Gonzalez-Olvera, J.J., Garza-Villarreal, E.A., 2019. Identifying cognitive deficits in cocaine dependence using standard tests and machine learning. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 95, 109709. <https://doi.org/10.1016/j.pnpbp.2019.109709>.
- Johnson, D.M., 2013. Forecasting corn and soybean yields in the United States utilizing pre- and within-season remotely sensed variables 2013, B54C-02.
- Jung, J., Maeda, M., Chang, A., Bhandari, M., Ashapure, A., Landivar-Bowles, J., 2021. The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. *Curr. Opin. Biotechnol. Food Biotechnol. Plant Biotechnol.* 70, 15–22. <https://doi.org/10.1016/j.cobio.2020.09.003>.
- Kalogiannidis, S., Kalfas, D., Chatzitheodoridis, F., Papaevangelou, O., 2022. Role of crop-protection technologies in sustainable agricultural productivity and management. *Land* 11, 1680. <https://doi.org/10.3390/land11101680>.
- Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., Shearer, S., 2018. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Comput. Electron. Agric.* 153, 213–225. <https://doi.org/10.1016/j.compag.2018.07.016>.
- Kim, D.-W., Yun, H.S., Jeong, S.-J., Kwon, Y.-S., Kim, S.-G., Lee, W.S., Kim, H.-J., 2018. Modeling and testing of growth status for chinese cabbage and white radish with UAV-based RGB imagery. *Remote Sens. (Basel)* 10, 563. <https://doi.org/10.3390/rs10040563>.
- Klemas, V.V., 2015. Coastal and environmental remote sensing from unmanned aerial vehicles: An overview. *J. Coast. Res.* 31, 1260–1267. <https://doi.org/10.2112/JCOASTRES-D-15-00005.1>.
- Kopitar, L., Kocabek, P., Cilar, L., Sheikh, A., Stiglic, G., 2020. Early detection of type 2 diabetes mellitus using machine learning-based prediction models. *Sci. Rep.* 10, 11981. <https://doi.org/10.1038/s41598-020-68771-z>.
- Kuhn, M., 2008. Building predictive models in R using the caret package. *J. Stat. Softw.* 28, 1–26. <https://doi.org/10.18637/jss.v028.i05>.
- Lakshminarayanan, K., Harp, S.A., Samad, T., 1999. Imputation of missing data in industrial databases. *Appl. Intell.* 11, 259–275. <https://doi.org/10.1023/A:1008334909089>.
- Lattanzi, D., Miller, G., 2017. Review of robotic infrastructure inspection systems. *J. Infrastruct. Syst.* 23, 04017004. [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000353](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000353).
- Li, H., Chen, L., Wang, Z., Yu, Z., 2019. Mapping of river terraces with low-cost UAV based structure-from-motion photogrammetry in a complex terrain setting. *Remote Sens. (Basel)* 11, 464. <https://doi.org/10.3390/rs11040464>.
- Li, Y., Liu, C., 2019. Applications of multirotor drone technologies in construction management. *Int. J. Constr. Manag.* 19, 401–412. <https://doi.org/10.1080/15623599.2018.1452101>.
- Li, W., Niu, Z., Chen, H., Li, D., Wu, M., Zhao, W., 2016. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecol. Ind.* 67, 637–648. <https://doi.org/10.1016/j.ecolind.2016.03.036>.
- Linaza, M.T., Posada, J., Bund, J., Eisert, P., Quartulli, M., Döllner, J., Pagani, A., G. Olaizola, I., Barriguinha, A., Moysiadis, T., Lucat, L., 2021. Data-Driven Artificial Intelligence Applications for Sustainable Precision Agriculture. *Agronomy* 11, 1227. <https://doi.org/10.3390/agronomy11061227>.
- Liu, Y., Feng, H., Yue, J., Fan, Y., Jin, X., Zhao, Y., Song, X., Long, H., Yang, G., 2022. Estimation of potato above-ground biomass using UAV-based hyperspectral images and machine-learning regression. *Remote Sens. (Basel)* 14, 5449. <https://doi.org/10.3390/rs14215449>.
- Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J.I., Benning, T.L., 2003. Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties.

- Agr. Ecosyst. Environ. 94, 205–220. [https://doi.org/10.1016/S0167-8809\(02\)00021-X](https://doi.org/10.1016/S0167-8809(02)00021-X).
- Löw, F., Biradar, C., Dubovyk, O., Fliemann, E., Akramkhanov, A., Narvaez Vallejo, A., Waldner, F., 2018. Regional-scale monitoring of cropland intensity and productivity with multi-source satellite image time series. *GISci. Remote Sens.* 55, 539–567. <https://doi.org/10.1080/15481603.2017.1414010>.
- Lussem, U., Bolten, A., Menne, J., Gnyp, M.L., Schellberg, J., Bareth, G., 2019a. Estimating biomass in temperate grassland with high resolution canopy surface models from UAV-based RGB images and vegetation indices. *JARS* 13, 034525. <https://doi.org/10.1111/1.JRS.13.034525>.
- Lussem, U., Bolten, A., Menne, J., Gnyp, M.L., Schellberg, J., Bareth, G., 2019b. Estimating biomass in temperate grassland with high resolution canopy surface models from UAV-based RGB images and vegetation indices. *JARS* 13, 034525. <https://doi.org/10.1111/1.JRS.13.034525>.
- Mahesh, B., 2019. Machine Learning Algorithms -A Review. *IJSR* 9 (1), 381–386. <https://doi.org/10.21275/ART20203995>.
- Maimaitijiang, M., Sagan, V., Sidike, P., Daloye, A.M., Erkbol, H., Fritsch, F.B., 2020. Crop monitoring using satellite/UAV data fusion and machine learning. *Remote Sens. (Basel)* 12, 1357. <https://doi.org/10.3390/rs12091357>.
- Marston, Z.P.D., Cira, T.M., Hodgson, E.W., Knight, J.F., Macrae, I.V., Koch, R.L., 2020. Detection of stress induced by Soybean Aphid (Hemiptera: Aphididae) using multispectral imagery from unmanned aerial vehicles. *J. Econ. Entomol.* 113, 779–786. <https://doi.org/10.1093/jee/toz306>.
- Maulud, D., Abdulazeez, A.M., 2020. A review on linear regression comprehensive in machine learning. *JASTI* 1, 140–147. <https://doi.org/10.38094/jasti1457>.
- Mazur, M., 2016. Six Ways Drones Are Revolutionizing Agriculture [WWW Document]. MIT Technology Review. URL <https://www.technologyreview.com/2016/07/20/158748/six-ways-drones-are-revolutionizing-agriculture/> (accessed 2.2.23).
- MicaSense, A.S., 2021. Atlas Flight by MicaSense, Inc. [WWW Document]. AppAdvice. URL [/app/atlas-flight/1103867349](http://app/atlas-flight/1103867349) (accessed 5.12.21).
- Mochida, K., Koda, S., Inoue, K., Hirayama, T., Tanaka, S., Nishii, R., Melgani, F., 2019. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective. *GigaScience* 8, giy153. <https://doi.org/10.1093/gigascience/giy153>.
- Mohammed, B., Awan, I., Ugail, H., Younas, M., 2019. Failure prediction using machine learning in a virtualised HPC system and application. *Cluster Comput.* 22, 471–485. <https://doi.org/10.1007/s10586-019-02917-1>.
- Moran, M.S., Inoue, Y., Barnes, E.M., 1997. Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sens. Environ.* 61, 319–346. [https://doi.org/10.1016/S0034-4257\(97\)00045-X](https://doi.org/10.1016/S0034-4257(97)00045-X).
- Nadkarni, P., 2016. Chapter 10 - Core Technologies: Data Mining and "Big Data," in: Nadkarni, P. (Ed.), *Clinical Research Computing*. Academic Press, pp. 187–204. <https://doi.org/10.1016/B978-0-12-803130-8.00010-5>.
- Nex, F., Remondino, F., 2014. UAV for 3D mapping applications: a review. *Appl. Geomat.* 6, 1–15. <https://doi.org/10.1007/s12518-013-0120-x>.
- Nickmilder, C., Tedde, A., Dufrasne, I., Lessire, F., Tychoon, B., Curnel, Y., Bindelle, J., Soyeurt, H., 2021. Development of machine learning models to predict compressed sward height in walloon pastures based on sentinel-1, sentinel-2 and meteorological data using multiple data transformations. *Remote Sens. (Basel)* 13, 408. <https://doi.org/10.3390/rs13030408>.
- Niu, Y., Zhang, L., Zhang, H., Han, W., Peng, X., 2019. Estimating above-ground biomass of maize using features derived from UAV-based RGB imagery. *Remote Sens. (Basel)* 11, 1261. <https://doi.org/10.3390/rs11111261>.
- Osborne, J., Waters, E., 2019. Four assumptions of multiple regression that researchers should always test. *Pract. Assess. Res. Eval.* 8 <https://doi.org/10.7275/r222-hv23>.
- Oscio, L.P., Junior, J.M., Ramos, A.P.M., Furuya, D.E.G., Santana, D.C., Teodoro, L.P.R., Gonçalves, W.N., Baio, F.H.R., Pistori, H., Junior, C.A. da S., Teodoro, P.E., 2020. Leaf Nitrogen Concentration and Plant Height Prediction for Maize Using UAV-Based Multispectral Imagery and Machine Learning Techniques. *Remote Sensing* 12, 3237. <https://doi.org/10.3390/rs12193237>.
- Oumar, Z., Mutanga, O., 2014. Integrating environmental variables and WorldView-2 image data to improve the prediction and mapping of Thaumastocoris peregrinus (bronze bug) damage in plantation forests. *ISPRS J. Photogramm. Remote Sens.* 87, 39–46. <https://doi.org/10.1016/j.isprsjprs.2013.10.010>.
- Panday, U.S., Shrestha, N., Maharjan, S., Pratihast, A., Shahnewaz, S., Shrestha, K., Aryal, J., 2020. Correlating the plant height of wheat with above-ground biomass and crop yield using drone imagery and crop surface model. A case study from Nepal. *Drones* 4, 28. <https://doi.org/10.3390/drones4030028>.
- Partel, V., Nunes, L., Stansly, P., Ampatzidis, Y., 2019. Automated vision-based system for monitoring Asian citrus psyllid in orchards utilizing artificial intelligence. *Comput. Electron. Agric.* 162, 328–336. <https://doi.org/10.1016/j.compag.2019.04.022>.
- Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., Athanasiadis, I.N., 2021. Machine learning for large-scale crop yield forecasting. *Agr. Syst.* 187, 103016 <https://doi.org/10.1016/j.agrsy.2020.103016>.
- Payero, J.O., Neale, C.M.U., Wright, J.L., 2004. Comparison of eleven vegetation indices for estimating plant height of alfalfa and grass. *Appl. Eng. Agric.* 20, 385–393. <https://doi.org/10.13031/2013.16057>.
- Pérez-Ortiz, M., Peña, J.M., Gutiérrez, P.A., Torres-Sánchez, J., Hervás-Martínez, C., López-Granados, F., 2015. A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method. *Appl. Soft Comput.* 37, 533–544. <https://doi.org/10.1016/j.asoc.2015.08.027>.
- Paul J. Pinter, Jr., Jerry L. Hatfield, James S. Schepers, Edward M. Barnes, M. Susan Moran, Craig S.T. Daughtry, Dan R. Upchurch, 2003. Remote Sensing for Crop Management. Photogrammetric engineering and remote sensing 69, 647–664. <https://doi.org/10.14358/PERS.69.6.647>.
- Poudyal, C., Sandhu, H., Ampatzidis, Y., Odero, D.C., Arbelo, O.C., Cherry, R.H., Costa, L.F., 2023. Prediction of morpho-physiological traits in sugarcane using aerial imagery and machine learning. *Smart Agric. Technol.* 3, 100104 <https://doi.org/10.1016/j.atech.2022.100104>.
- Pranga, J., Borra-Serrano, I., Aper, J., De Swaeef, T., Ghesquiere, A., Quataert, P., Roldán-Ruiz, I., Janssens, I.A., Ruyschaert, G., Lootens, P., 2021. Improving accuracy of herbage yield predictions in perennial ryegrass with UAV-based structural and spectral data fusion and machine learning. *Remote Sens. (Basel)* 13, 3459. <https://doi.org/10.3390/rs13173459>.
- Prasad, A.K., Chai, L., Singh, R.P., Kafatos, M., 2006. Crop yield estimation model for Iowa using remote sensing and surface parameters. *Int. J. Appl. Earth Obs. Geoinf.* 8, 26–33. <https://doi.org/10.1016/j.jag.2005.06.002>.
- Probst, P., Wright, M.N., Boulesteix, A.-L., 2019. Hyperparameters and tuning strategies for random forest. *WIREs Data Min. Knowl. Discovery* 9, e1301.
- Rajaveni, S.P., Brindha, K., Elango, L., 2017. Geological and geomorphological controls on groundwater occurrence in a hard rock region. *Appl. Water Sci.* 7, 1377–1389. <https://doi.org/10.1007/s13201-015-0327-6>.
- Rashidi, H.H., Tran, N.K., Betts, E.V., Howell, L.P., Green, R., 2019. Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. *Academic Pathology* 6, 2374289519873088. <https://doi.org/10.1177/2374289519873088>.
- Rasmussen, J., Ntakos, G., Nielsen, J., Svensgaard, J., Poulsen, R., Christensen, S., 2016. Are vegetation indices derived from consumer-grade cameras mounted on UAVs sufficiently reliable for assessing experimental plots? *Eur. J. Agron.* 74, 75–92. <https://doi.org/10.1016/j.eja.2015.11.026>.
- Reedha, R., Deriquebourg, E., Canals, R., Hafiane, A., 2022. Transformer neural network for weed and crop classification of high resolution UAV images. *Remote Sens. (Basel)* 14, 592. <https://doi.org/10.3390/rs14030592>.
- Ren, X., Mi, Z., Georgopoulos, P.G., 2020. Comparison of machine learning and land use regression for fine scale spatiotemporal estimation of ambient air pollution: Modeling ozone concentrations across the contiguous United States. *Environ. Int.* 142, 105827 <https://doi.org/10.1016/j.envint.2020.105827>.
- Revenga, J.C., Trepekli, K., Oehmcke, S., Jensen, R., Li, L., Igel, C., Gieseke, F.C., Friberg, T., 2022. Above-ground biomass prediction for croplands at a sub-meter resolution using UAV-LiDAR and machine learning methods. *Remote Sens. (Basel)* 14, 3912. <https://doi.org/10.3390/rs14163912>.
- Robles, W., Madsen, J.D., Wersal, R.M., 2010. Potential for remote sensing to detect and predict herbicide injury on Waterhyacinth (*Eichhornia crassipes*). *Invasive Plant Sci. Manage.* 3, 440–450. <https://doi.org/10.1614/IPSM-D-09-00040.1>.
- Rosser, J.C., Vignesh, V., Terwilliger, B.A., Parker, B.C., 2018. Surgical and medical applications of drones: A comprehensive review. *JLSL* 22 (e2018), 00018. <https://doi.org/10.4293/JSLS.2018.00018>.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS.
- Ruengvirayudh, P., Brooks, G., 2016. Comparing stepwise regression models to the best-subsets models, or, the art of stepwise. *General Linear Model J.*
- Samseemoung, G., Soni, P., Jayasuriya, H.P.W., Salokhe, V.M., 2012. Application of low altitude remote sensing (LARS) platform for monitoring crop growth and weed infestation in a soybean plantation. *Precis. Agric.* 13, 611–627. <https://doi.org/10.1007/s11119-012-9271-8>.
- Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., Brenning, A., 2019. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecol. Model.* 406, 109–120. <https://doi.org/10.1016/j.ecolmodel.2019.06.002>.
- Selvaraj, M.G., Valderrama, M., Guzman, D., Valencia, M., Ruiz, H., Achcarjee, A., 2020. Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground traits in cassava (*Manihot esculenta Crantz*). *Plant Methods* 16, 87. <https://doi.org/10.1186/s13007-020-00625-1>.
- Shakoor, N., Lee, S., Mockler, T.C., 2017. High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Current Opinion in Plant Biology*, 38 Biotic interactions 2017 38, 184–192. <https://doi.org/10.1016/j.jpb.2017.05.006>.
- Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. [https://doi.org/10.1016/S0034-4257\(02\)00010-X](https://doi.org/10.1016/S0034-4257(02)00010-X).
- Singh, K.K., Frazier, A.E., 2018. A meta-analysis and review of unmanned aircraft system (UAS) imagery for terrestrial applications. *Int. J. Remote Sens.* 39, 5078–5098. <https://doi.org/10.1080/01431161.2017.1420941>.
- Soltis, P.S., Nelson, G., Zare, A., Meineke, E.K., 2020. Plants meet machines: Prospects in machine learning for plant biology. *Appl. Plant Sci.* 8, e11371.
- Song, Y., Wang, J., 2019. Winter wheat canopy height extraction from UAV-based point cloud data with a moving cuboid filter. *Remote Sens. (Basel)* 11, 1239. <https://doi.org/10.3390/rs11101239>.
- Strobl, C., Malley, J., Tutz, G., 2009. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol. Methods* 14, 323–348. <https://doi.org/10.1037/a0016973>.
- Sun, S., Wang, C., Ding, H., Zou, Q., 2020. Machine learning and its applications in plant molecular studies. *Brief. Funct. Genomics* 19, 40–48. <https://doi.org/10.1093/bfgp/elz036>.
- Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P., 2003. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* 43, 1947–1958. <https://doi.org/10.1021/ci034160g>.

- Teodoro, P.E., Teodoro, L.P.R., Baio, F.H.R., da Silva Junior, C.A., dos Santos, R.G., Ramos, A.P.M., Pinheiro, M.M.F., Osco, L.P., Gonçalves, W.N., Carneiro, A.M., Junior, J.M., Pistori, H., Shiratsuchi, L.S., 2021. Predicting days to maturity, plant height, and grain yield in soybean: A machine and deep learning approach using multispectral data. *Remote Sens. (Basel)* 13, 4632. <https://doi.org/10.3390/rs13224632>.
- Thelen, K.D., Kravchenko, A.N., Lee, C.D., 2004. Use of optical remote sensing for detecting herbicide injury in soybean. *Weed Technol.* 18, 292–297. <https://doi.org/10.1614/WT-03-049R2>.
- Thenkabail, P.S., Ward, A.D., Lyon, J., Merry, C.J., 1994. Thematic mapper vegetation indices for determining soybean and corn growth parameters. *PE&RS*, 60 (4), 437–442.
- Tilly, N., Aasen, H., Bareth, G., 2015. Fusion of plant height and vegetation indices for the estimation of barley biomass. *Remote Sens. (Basel)* 7, 11449–11480. <https://doi.org/10.3390/rs70911449>.
- van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* 177, 105709. <https://doi.org/10.1016/j.compag.2020.105709>.
- Veysi, S., Naseri, A.A., Hamzeh, S., Bartholomeus, H., 2017. A satellite based crop water stress index for irrigation scheduling in sugarcane fields. *Agric. Water Manag.* 189, 70–86.
- Vijayakumar, V., Ampatzidis, Y., Costa, L., 2023. Tree-level citrus yield prediction utilizing ground and aerial machine vision and machine learning. *Smart Agric. Technol.* 3, 100077. <https://doi.org/10.1016/j.atech.2022.100077>.
- Viijanen, N., Honkavaara, E., Näsi, R., Hakala, T., Niemeläinen, O., Kaivosoja, J., 2018. A novel machine learning method for estimating biomass of grass swards using a photogrammetric canopy height model, images and vegetation indices captured by a drone. *Agriculture* 8, 70. <https://doi.org/10.3390/agriculture8050070>.
- Wang, Y., Shi, X., Lei, L., Fung, J.-C.-H., 2022. Deep learning augmented data assimilation: Reconstructing missing information with convolutional autoencoders. *Mon. Weather Rev.* 150, 1977–1991. <https://doi.org/10.1175/MWR-D-21-0288.1>.
- Wang, L., Zhou, X., Zhu, X., Dong, Z., Guo, W., 2016. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J.* 4, 212–219. <https://doi.org/10.1016/j.cj.2016.01.008>.
- Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiyama-Kanegae, H., Kobayashi, M., Yano, K., Tokunaga, T., Fujiwara, T., Tsutsumi, N., Iwata, H., 2017. High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling. *Front. Plant Sci.* 8, 421. <https://doi.org/10.3389/fpls.2017.00421>.
- Whitmire, C.D., Vance, J.M., Rasheed, H.K., Missaoui, A., Rasheed, K.M., Maier, F.W., 2021. Using Machine Learning and Feature Selection for Alfalfa Yield Prediction. *AI* 2, 71–88. <https://doi.org/10.3390/ai2010006>.
- Wiseman, G., McNairn, H., Homayouni, S., Shang, J., 2014. RADARSAT-2 polarimetric SAR response to crop biomass for agricultural production monitoring. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 4461–4471. <https://doi.org/10.1109/JSTARS.2014.2322311>.
- Xiang, T.-Z., Xia, G.-S., Zhang, L., 2019. Mini-unmanned aerial vehicle-based remote sensing: Techniques, applications, and prospects. *IEEE Geosci. Remote Sens. Mag.* 7, 29–63. <https://doi.org/10.1109/MGRS.2019.2918840>.
- Xu, Y., Ghamisi, P., 2022. Consistency-regularized region-growing network for semantic segmentation of urban scenes with point-level annotations. *IEEE Trans. Image Process.* 31, 5038–5051. <https://doi.org/10.1109/TIP.2022.3189825>.
- Yalcin, H., 2017. Plant phenology recognition using deep learning: Deep-Pheno, in: 2017 6th International Conference on Agro-Geoinformatics. Presented at the 2017 6th International Conference on Agro-Geoinformatics, pp. 1–5. <https://doi.org/10.1109/Agro-Geoinformatics.2017.8046996>.
- Yao, H., Qin, R., Chen, X., 2019. Unmanned aerial vehicle for remote sensing applications—A review. *Remote Sens. (Basel)* 11, 1443. <https://doi.org/10.3390/rs11121443>.
- Yu, W., Liu, T., Valdez, R., Gwinn, M., Khoury, M.J., 2010. Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *BMC Med. Inform. Decis. Mak.* 10, 16. <https://doi.org/10.1186/1472-6947-10-16>.
- Yue, J., Feng, H., Yang, G., Li, Z., 2018. A comparison of regression techniques for estimation of above-ground winter wheat biomass using near-surface spectroscopy. *Remote Sens. (Basel)* 10, 66. <https://doi.org/10.3390/rs10010066>.
- Zhang, C., Kovacs, J.M., 2012. The application of small unmanned aerial systems for precision agriculture: a review. *Precis. Agric.* 13, 693–712. <https://doi.org/10.1007/s11119-012-9274-5>.
- Zhang, Y., Li, M., Li, G., Li, J., Zheng, L., Zhang, M., Wang, M., 2022. Multi-phenotypic parameters extraction and biomass estimation for lettuce based on point clouds. *Measurement* 204, 112094. <https://doi.org/10.1016/j.measurement.2022.112094>.
- Zhang, L., Niu, Y., Zhang, H., Han, W., Li, G., Tang, J., Peng, X., 2019. Maize canopy temperature extracted from UAV thermal and RGB imagery and its application in water stress monitoring. *Front. Plant Sci.* 10, 1270. <https://doi.org/10.3389/fpls.2019.01270>.
- Zheng, C., Abd-Elrahman, A., Whitaker, V., Dalid, C., 2022. Prediction of strawberry dry biomass from UAV multispectral imagery using multiple machine learning methods. *Remote Sens. (Basel)* 14, 4511. <https://doi.org/10.3390/rs14184511>.
- Zou, K.H., Tuncali, K., Silverman, S.G., 2003. Correlation and simple linear regression. *Radiology* 227, 617–628. <https://doi.org/10.1148/radiol.2273011499>.