The International Society of Precision Agriculture presents the
15th International Conference on Precision Agriculture
26–29 JUNE 2022
Minneapolis Marriott City Center | Minneapolis, Minnesota USA

Hay yield estimation using UAV-based imagery with machine learning technology

Kyuho Lee[A], Kenneth A. Sudduth[B], Jianfeng Zhou[C]

A: Department of Biomedical, Biological and Chemical Engineering, University of Missouri, Columbia, Missouri 65211, USA

B: USDA-ARS Cropping Systems and Water Quality Research Unit, Columbia, Missouri 65211, USA

C: Division of Plant Science and Technology, University of Missouri, Columbia, Missouri 65211, USA

A paper from the Proceedings of the
15th International Conference on Precision Agriculture
June 26-29, 2022
Minneapolis, Minnesota, United States

Abstract. *Yield monitoring systems are widely used commercially in grain crops to map yields at a scale of a few meters. However, such high-resolution yield monitoring and mapping for hay and forage crops has not been commercialized. Most commercial hay yield monitoring systems only obtain the weight of individual bales, making it difficult to map and understand the spatial variability in hay yield. This study investigated the feasibility of an unmanned aerial vehicle (UAV)-based remote sensing system for the estimation and mapping of hay yield by machine learning models. Data were obtained during harvest of a 35-ha hay field with mixture of red clover and timothy grass in June of 2021. A RGB camera consisting of three bands (red, blue, and green) attached to a UAV was used to acquire images at a flight height of 20 m For calibration, 110 ground truth hay yield measurements were collected from 1 m^2 quadrats. Image features, such as color space components, vegetation indices and texture features, the proportion of grass in samples, and moisture content of samples were extracted from the images or ground truth samples, and were used to estimate the hay mass yield. For yield estimation, a simple random forest machine learning model was trained and tested with the stratified random sampling method using a split ratio of 70:30. Using the recursive feature elimination algorithm, we selected explanatory features for use in the random forest regression model. The most accurate model estimated hay wet mass with r^2 = 0.79, RMSE = 251.05 g/m^2, and MAE = 187.59 g/m^2). The results of this research provide information to aid in selection of an appropriate analysis method for hay estimation using UAV imagery. In future research, the models developed here will be applied to whole-field imagery for creating hay yield maps.*

Keywords. *Yield monitoring, UAV, Remote sensing, Machine learning, Hay yield*

Introduction

Forage and hay production is essential to provide feed to animals such as cattle and horses. Commercial systems to monitor and estimate hay yield are based on the weight of bales, using load cells or scales (Maguire et al., 2003). Different types of sensors such as mass-flow sensors have been considered to improve yield estimation accuracy. Previous studies have tried to find appropriate yield monitoring solutions to be installed on hay balers. Maughan et al. (2012) summarized research and development of sensors and systems designed to collect mass flow and yield data of forage and hay. Various hay yield estimation systems have been developed and tested in field conditions, including: 1) feedroll displacement, 2) crop impact force, 3) torque sensor, and 4) others (e.g., load cell and strain gauges). The accuracy of these yield estimation systems could be as high as $R^2 > 0.8$ under ideal conditions. However, several factors, such as crop toughness and low plant density could decrease the prediction accuracies, and Maughan et al. (2012) suggested that additional sensing technologies should be considered in the future. Despite high accuracy, those systems are not able to monitor crops temporally and spatially because the data are only collected on the day of harvest.

Remote sensing technology detects and provides information about crops by equipping satellites, manned aircraft, or unmanned aerial vehicles (UAV) with a variety of sensors (Rueda-Ayala et al., 2019). In the agriculture field, remote sensing can temporally and spatially monitor crops (e.g., biomass and nitrogen content; Näsi et al., 2018). For example, temporal parameters from crop images can help improve vegetation models that estimate plant growth, disease, and status (Senf et al., 2017; Rueda-Ayala et al., 2019). Although satellite images can be useful in monitoring and detecting vegetation at a regional scale efficiently, the collected data are low in resolution, are costly, and may have irregular and infrequent acquisitions (e.g., weekly or bi-weekly), which limit their use for within-field management. With advances in UAVs and emerging image processing and analysis technologies, UAV-based remote sensing systems have become popular due to high resolution, ability to control acquisitions, and relatively low cost.

Agricultural remote sensing technology results in big data, primarily through the spatial component facilitated by global positioning (GPS) and geographic information systems (GIS; Huang et al., 2018). For this reason, successful precision agriculture applications depend on managing, processing, and analyzing big data. Machine learning through a deep architecture can analyze both labeled and unlabeled data, so it is suitable to be applied to big data analysis (Liu, 2015). For instance, random forest (RF) and support vector machine (SVM) algorithms have been widely used for classification and regression in remote sensing technology (Abdulridha et al., 2019; Feng et al., 2019; Osco et al., 2020). Notably, the deep learning technique, which can extract features automatically, has been specialized to analyze and model data through deeper neural networks (Ghamisi et al., 2017).

In previous research estimating biomass, models based on convolutional neural networks (CNNs) were used. According to LeCun et al. (2015), CNNs use multiple layer neural networks equipped with convolutional and pooling layers, and they have a strong ability to recognize complicated features. To predict forage yield of different genotypes, Castro et al. (2020) used two CNN models (AlexNet with 8 layers and ResNet18 with 18 layers) and compared them with the VGCNet model (11 layers), which they used in previous research. In another case, AlexNet (8 layers), ResNeXt50 (50 layers), DarkNet53 (53 layers), MaCNN (5 layers), and LF-CNN (10 layers) architectures were used with RGB images to estimate guineagrass dry matter yield (de Oliveira et al., 2021).

In previous research estimating biomass by remote sensing, multiple image acquisitions obtained over the growing season were often used. For example, Kattenborn et al. (2021) used weekly or bi-weekly data collection over 200 plots that were 45 m^2 in size for creating training and test datasets. Although this may be needed for best results, the amount of effort needed could be more than would be feasible in commercial yield mapping. An alternative approach using a single imaging date near harvest and requiring less ground truth data collection would be more attractive

to producers. Therefore, the goal of this study was to evaluate the performance of a UAV-based remote sensing system in predicting hay yield with a single harvest-time image collection. Specific objectives included 1) applying a random forest machine learning model to quantify hay yield based on RGB images and ground truth data obtained for 1m x 1m quadrats, 2) and evaluating the usefulness of a range of image features, texture features, and vegetation indices in the model.

Materials and Methods

Data collection based on UAV system

The field research was conducted at the Central Mississippi River Basin (CMRB) site of the USDA's Long-Term Agroecosystem Research (LTAR) network (39˚13 N, 92˚07 W), near Centralia, Missouri (Sadler et al., 2016). A 35-ha research field in a corn-soybean-wheat-hay rotation, was seeded with a mixture of red clover (*Trifolium pratense* L) and timothy grass (*Phleum pratense*) in March 2020, in conjunction with an early-spring nitrogen fertilization. The hay was managed according to standard management protocols in a rainfed production system. At the time of harvest in June 2021, some portions of the field were dominated by timothy, while others were predominantly red clover (Fig. 1).

Fig 1. The hay crop in the experimental field was a mixture of timothy grass and red clover.

Data collection included ground-truth biomass data at selected sample locations and UAV-based remote sensing data across the field (Fig 2). All field data were collected at the maturity stage (e.g., reproductive and flowering) of hay on June 4th and 9th 2021, within one day of harvest. Sampling sites, each containing a variable mixture of timothy and red clover were established in two areas – along the west (roadside) edge of the field (June 4th data collection) and on an east-west transect in the northern part of the field (June 9th data collection). Sixty sampling sites were identified at the west edge, and fifty at the northern area. On the west edge, data collection was divided into two parts due to flight time limitations of the UAV: the northern (11:00 am to 12:00 pm) and southern (12:00 pm to 13:00 pm) part.

Each sampling site was identified using a 1 m × 1 m quadrat made of 2 cm diameter PVC pipe. The location of each quadrat was measured using an RTK GPS system (Reach RS+, Emlid, Saint Petersburg, Russia) at the center of the quadrats. The height of the vegetation mixture was measured at ten locations inside each quadrat using a stick ruler with 1 cm precision. Then, all the vegetation within the quadrat was cut at approximately 2 cm above the ground using a handheld grass cutter (STIHL FS 90R, STIHL Corporation, Waiblingen, Germany) after the sensing data were collected. All vegetation from each quadrat was collected into a bag that was sealed to reduce the moisture loss and labeled properly.

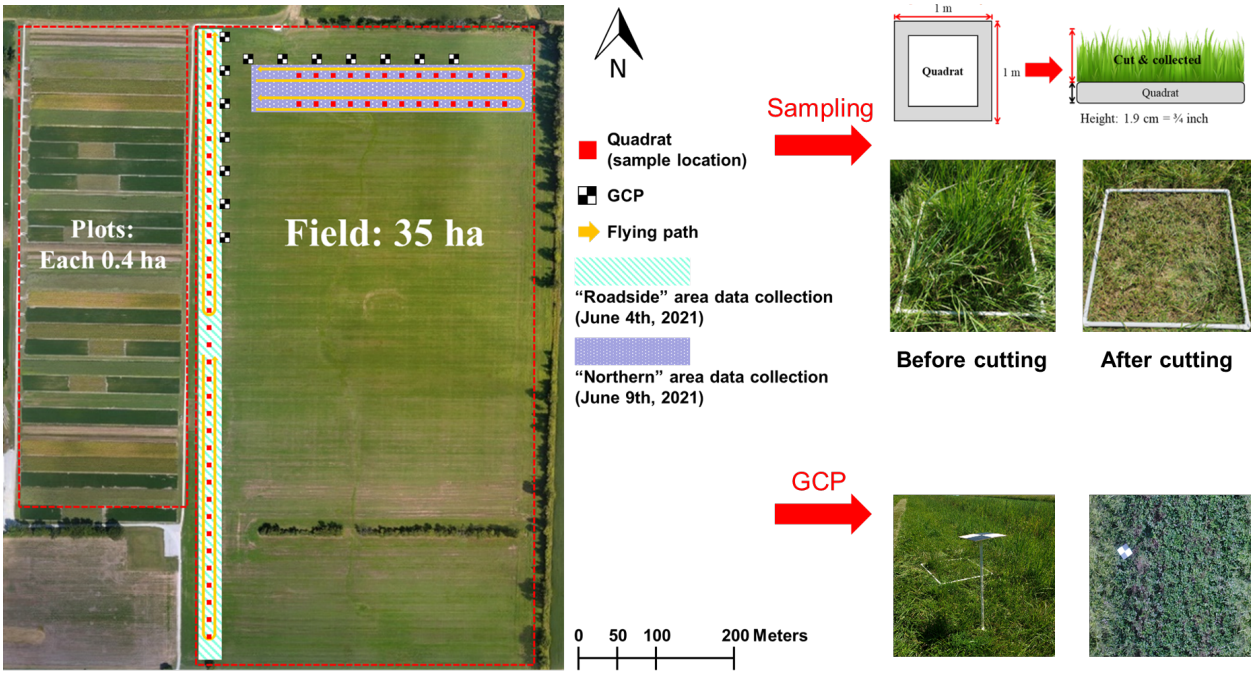

Fig 2. Information of experiment (left: location of data collection in the field – "Roadside" and "Northern" area, right: quadrat, sampling and ground control points (GCPs) images).

Vegetation samples were processed on the day of data collection. First, the overall sample wet mass was measured and then a subsample was obtained and its wet mass measured. The subsamples were dried at 105 ℃ for 24 hours using a laboratory oven (SHEL LAB SMO28-2, Sheldon Manufacturing Inc., Cornelius, OR, USA) to determine subsample dry mass. Then, the moisture content of the samples and dry sample mass were calculated by Eqs. 1 and 2. The proportion of grass and clover in the hay varied across the field. A visual rating of that proportion was obtained at each sampling location using a 1-5 rating as shown in Table 1.

$$\text{Mositure content (dry basis, \%)} = \frac{(Subsample\ wet\ mass - subsample\ dry\ mass)}{(Subsample\ dry\ weight)} \times 100 \quad (1)$$

$$\text{Dry mass (g)} = \frac{Subsample\ dry\ mass}{Subsample\ wet\ mass} \times Sample\ wet\ mass \quad (2)$$

Table 1. The proportion of grass in samples

Group number	Proportion
1	Mostly timothy grass
2	Approximately 2:1 timothy to red clover
3	Equal
4	Approximately 1:2 timothy to red clover
5	Mostly red clover

Fig 3 shows the workflow of the study. The left column of Fig 3 shows the division of the procedure into five parts, and the specific tasks for each part are shown to the right.

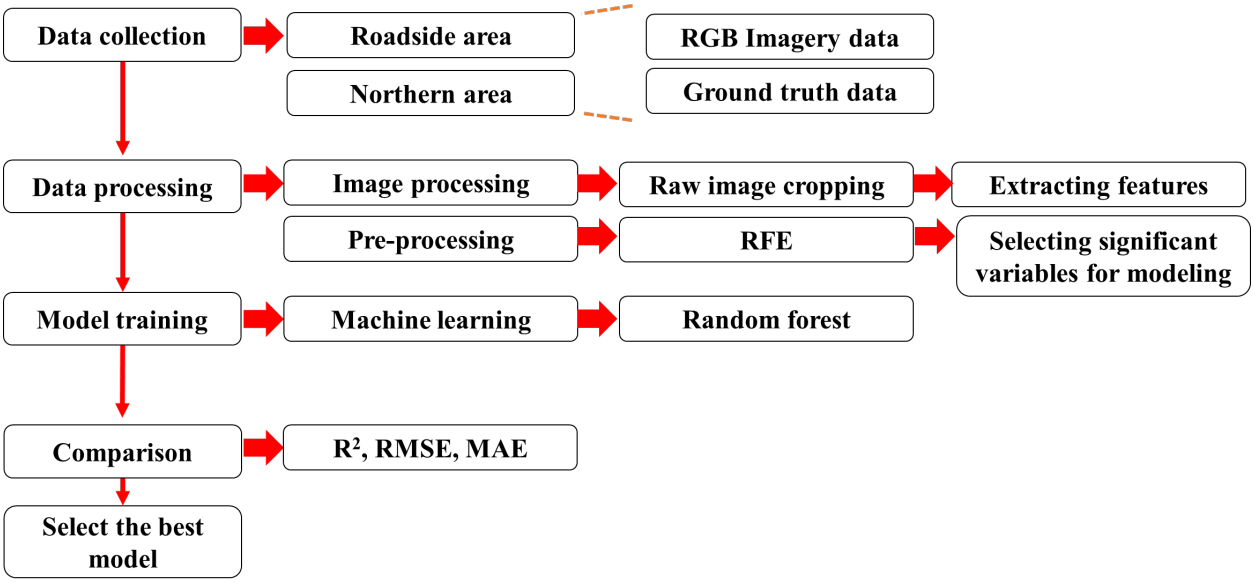

Fig 3. Workflow of the study divided into five parts: data collection, data processing, model training, comparison, and selecting the best model among the results

UAV system

A UAV remote imaging system (DJI Phantom 4 Advanced, DJI, Shengzhen, Guangdong, China) equipped with a RGB camera was used to acquire images at 20 m above ground level (AGL) resulting in a ground sampling distance (GSD) of 5mm/pixel. The RGB camera was configured to take time-lapse images at 0.5 frame per second (fps) using a UAV control software (Auto Pilot, Hangar Technology, Austin, TX, USA) installed on an iPad mini 4 (Apple Inc., Cupertino, CA, US).

The UAV platform was controlled using the flight control app Litchi (VC Technology Ltd, London, UK) for RGB image sensing. Images were acquired at 10 am to 1 pm Central Daylight Time on June 4th, 2021 for the roadside, and during the same time interval on June 9th, 2021 for the northern area. Both days had a clear sky with occasional strong wind. Images were taken at 20 m AGL, at a flight speed of 7 km/h, following a zigzag path to cover the field with forward overlap ≥ 70% and side overlap ≥ 65%.

Image processing

Fig 4 shows the specific image processing procedures used in the study. First, we stitched RGB images using the UAV image processing software Agisoft Metashape (Agisoft LLC, St. Petersburg, Russia) to create orthomosaic images and digital elevation models (DEM) for further processing. However, unexpected weather factors (e.g., strong wind) affected the UAV and RGB camera on June 4th around 11 am to 12 pm, causing rapid UAV elevation changes that resulted in poor quality of the stitched image. Therefore, to obtain higher quality data, the raw images were used to extract imagery information and train machine learning models. The raw images corresponding to the quadrats were located, rotated to align with the quadrat, and then cropped to the area enclosed by the qudrat. Then, Otsu's segmentation method (Otsu, 1979) was applied to the cropped images for segmentation of hay from images to reduce the effects of soil background. The cropped and segmented images were used to calculate pixel values of vegetation indices based on RGB channels for each quadrat using Matlab (R2019b, MathWorks, Natick, MA, US).

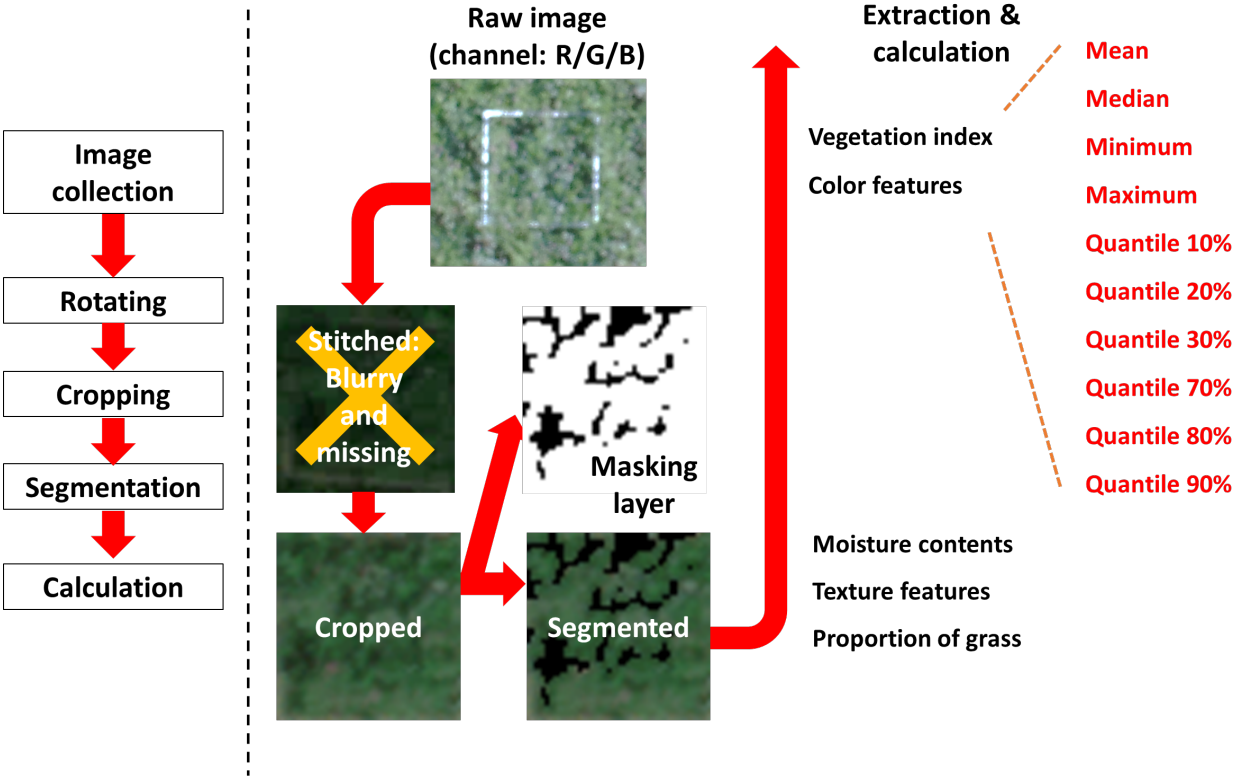

Fig 4. Image processing procedure from image collection to extraction of features.

Summary statistics of the individual pixel values from each quadrat were calculated to condense the information for modeling. Ten different statistics were calculated for each vegetation index and color feature, and models using the different statistics were created and compared. The statistics calculated were: maximum, minimum, mean, median, and six different quantile values: 10%, 20%, 30%, 70%, 80%, and 90%. As a result, ten statistical values were obtained for each vegetation index and color feature within each quadrat.

Image feature extraction

Models were developed with either the dry or wet mass of the quadrat samples as the dependent variable. Independent variables used for training the models consisted of five groups: color features, vegetation indices, texture features, proportion of grass and clover in each sample and moisture content of each sample. Twelve color features were included to train the model: HSV for color space (hue, saturation, value), L*a*b for CIE Lab color space (L: lightness, a: green-red component, b: blue-yellow component), and YCbCr for YCbCr color space (Y: lima component, Cb: Blue-difference chroma component, Cr: Red-difference chroma component). Additionally, thirty vegetation indices based on RGB channels were included. Six texture features (contrast, correlation, energy, entropy, homogeneity, and variance) were calculated from a gray-level co-occurrence matrix (GLCM). These texture features were used to quantify the spatial relationship between adjacent or neighboring pixels within the quadrats (Haralick et al., 1973). Finally, to predict more specific hay dry mass, in a model using dry mass as a dependent variable we added moisture contents of samples as an independent variable. The color features, texture features, and vegetation indices are detailed in Table 2.

Table 2. Potential independent variables (color features, texture features, and vegetation indices)

Index	Equation	Feature type & reference
H (Hue)	$H = \begin{cases} \frac{60 \times (G-B)}{V - \min(R,G,B)} & if\ V = R \\ 120 + \frac{60 \times (B-R)}{V - \min(R,G,B)} & if\ V = G \\ 240 + \frac{60 \times (R-G)}{V - \min(R,G,B)} & if\ V = b \end{cases}$	Color space feature (HSV)
S (Saturation)	$S = \begin{cases} \frac{V - \min(R,G,B)}{V} & if\ V \neq 0 \\ 0 & if\ V = 0 \end{cases}$	Color space feature (HSV)
V (Value)	$V = \max(R,G,B)$	Color space feature (HSV)
L (Lightness)	No single equation	Color space feature (CIE L*a*b)
a (Green-Red component)	No single equation	Color space feature (CIE L*a*b)
b (Blue-Yellow component)	No single equation	Color space feature (CIE L*a*b)
Y	$Y = 16 + \frac{65.738 \times R}{256} + \frac{129.057 \times G}{256} + \frac{25.064 \times B}{256}$	Color space feature (YCrCb)
Cr	$C_r = 128 - \frac{112.439 \times R}{256} - \frac{94.154 \times G}{256} - \frac{18.285 \times B}{256}$	Color space feature (YCrCb)
Cb	$C_b = 128 - \frac{37.945 \times R}{256} - \frac{74.494 \times G}{256} + \frac{112.439 \times B}{256}$	Color space feature (YCrCb)
Contrast	$\sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} (i-j)^2 \cdot g^2(i,j)$	Texture feature
Correlation	$\sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} (i-\mu) \cdot (j-\mu) \cdot g(i,j)/\sigma^2$	Texture feature
Energy	$\sqrt{\sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} g^2(i,j)}$	Texture feature
Entropy	$\sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} g^2(i,j) \cdot \log(g(i,j))$	Texture feature
Homogeneity	$\sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} \frac{1}{1+(i-j)^2} \cdot g(i,j)$	Texture feature
Variance	$\sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} (i-\mu)^2 \cdot g(i,j)$	Texture feature
BCC (Blue Chromatic Coordinate Index)	$\frac{B}{(R+G+B)}$	Woebbecke et al., 1995
B-G (Blue-Green Difference)	$B - G$	De Swaef et al., 2021
BRVI (Blue Red Vegetation Index)	$\frac{(B-R)}{(B+R)}$	De Swaef et al., 2021
CI (Coloration Index)	$\frac{(R-B)}{R}$	Escadafal et al., 1994
CIVE (Color Index of Vegetation)	$0.441 \times R - 0.881 \times G + 0.385 \times B + 18.787$	Lee et al., 2021
ExG (Excess Green Index)	$2 \times G - B - R$	Woebbecke et al., 1995

ExG2 (Excess Green Index 2)	$\frac{(2 \times G - B - R)}{(R + G + B)}$	Woebbecke et al., 1995
ExGR (Excess Green minus Excess Red Index)	$\frac{(2 \times G - B - R)}{(R + G + B)} - \frac{(1.4 \times R - G)}{(R + G + B)}$	Meyer et al., 2004
ExR (Excess Red Index)	$\frac{(2 \times G - B - R)}{(R + G + B)} - \frac{(1.4 \times R - G)}{(R + G + B)}$	Meyer et al., 1999
G/R (Green Red Ratio)	$\frac{G}{R}$	Steele et al., 2009
GBVI (Green Blue Vegetation Index)	$\frac{(G - B)}{(G + B)}$	De Swaef et al., 2021
GCC (Green Chromatic Coordinate Index)	$\frac{G}{(R + G + B)}$	Gitelson et al., 2002
GLI (Green Leaf Index)	$\frac{(2 \times G - R - B)}{(2 \times G + R + B)}$	Baroni et al., 2004
G-R (Green Red Difference)	$G - R$	De Swaef et al., 2021
GRVI (Green Red Vegetation Index)	$\frac{(G - R)}{(G + R)}$	Gitelson et al., 2002
I (Intensity)	$\frac{(R + G + B)}{30.5}$	Escadafal et al., 1994
IF (Shape Index)	$\frac{(2R - G - B)}{(G - B)}$	Escadafal et al., 1994
IO (Simple Ratio Red/Blue Iron Oxide)	$\frac{R}{B}$	Hewson et al., 2001
IRG (Red Green Ratio Index)	$R - G$	Jacobsen et al., 1995
MGRVI (Modified Green Red Vegetation)	$\frac{(G^2 - R^2)}{(G^2 + R^2)}$	Bendig et al., 2015
NGRDI (Normalized Green Red Difference Index)	$\frac{(G - R)}{(G + R)}$	Gitelson et al., 2002
RCC (Red Chromatic Coordinate Index)	$\frac{R}{(R + G + B)}$	Woebbecke et al., 1995
RGR (Simple Ratio Red/Green Red-Green Ratio)	$\frac{R}{G}$	Gamon and Surfus, 1999
RGBVI (Red Green Blue Vegetation Index)	$\frac{(G^2 - B \times R)}{(G^2 + B \times R)}$	Bendig et al., 2015
RI (Redness Index)	$\frac{(R - G)}{(R + G)}$	Escadafal et al., 1994
TGI (Triangular Greeness Index)	$(-0.5) \times \{0.19 \times (R - G) - 0.12 \times (R - B)\}$	Hunt et al., 2011
VARI (Visible Atmosphercally Resistant Index)	$\frac{(G - R)}{(G + R - B)}$	Gitelson et al., 2002
VDVI (Visible-band Difference Vegetation Index)	$\frac{(2 \times G - R - B)}{(2 \times G + R + B)}$	Hu and Li, 2019
VEG (Vegetative Index)	$\frac{G}{(R^{0.667} \times B^{0.334})}$	Torres-Sánchez et al., 2014
WI (Woebbecke Index)	$\frac{(G - B)}{(R - G)}$	Woebbecke et al., 1995

Pre-processing for modeling

With the completion of the above image processing and feature calculation, pre-processing for training machine learning models was needed to avoid overfitted models. In this study, modeling and finding the optimal subset of the candidate predictor variables was done with recursive feature elimination (RFE). The RFE method is a feature selection technique that ranks features by importance and removes the least important or correlated variables from the models. The RFE uses a random forest or multiple linear regression algorithm to test combinations of features and ranks features by correlation and RMSE value. Finally, the algorithm removes the lower performing features and recommends those scored as having a high accuracy, from minimizing and optimizing a loss function (Huang et al., 2018).

We used Rstudio (RStudio version 4.4, RStudio, Boston, MA, US) with the "caret" package (version 6.0-92) that includes the RFE algorithm based on the Gini criterion. For training the RFE models, we input the above independent variables (Table 2) and dry mass of the quadrat samples as the dependent variables. Then, the algorithm randomly selected variables to evaluate the relationship using a random forest algorithm to test combinations of features. This process was carried out nine more times (10-fold cross validation). Finally, based on the mean values of statistical results from the cross-validation, the RFE algorithm recommended the optimal number of variables and significant variables to use for modeling as output (Fig 5).

Machine learning models

Random forest regression was used to predict hay mass. The "Random Forest" package (version 4.7-1) of RStudio was used in the process. To train and test models, the dataset was divided into 70% training and 30% test samples using stratified random sampling. Then, we used the significant number of variables and the significant variables from the RFE recommendation. When setting parameters for the random forest models, *mtry* was dependent on the number of recommended variables, and *ntree* was set to 1000.

Results and Discussion

Grass:clover mixture

Table 3 shows that all grass:clover mixture proportions were well-represented in the calibration samples, Red clover was more dominant in the samples from the western roadside area, while the proportions of grass and clover were similar in the northern calibration samples. The grass:clover proportion was related to the manually measured height of vegetation inside the quadrat (Fig. 6). Although there was considerable variation in height within a single classification, there was an obvious trend toward a lower median height with and increasing amount of clover.

Table 3. Classification based on the proportion of grass vs. clover vegetation inside quadrats.

Group number	Proportion	Roadside frequency	Northern frequency
1	Mostly timothy grass	10	17
2	Approximately 2:1 timothy to red clover	5	8
3	Equal	9	5
4	Approximately 1:2 timothy to red clover	17	9
5	Mostly red clover	19	11
Total		60	50

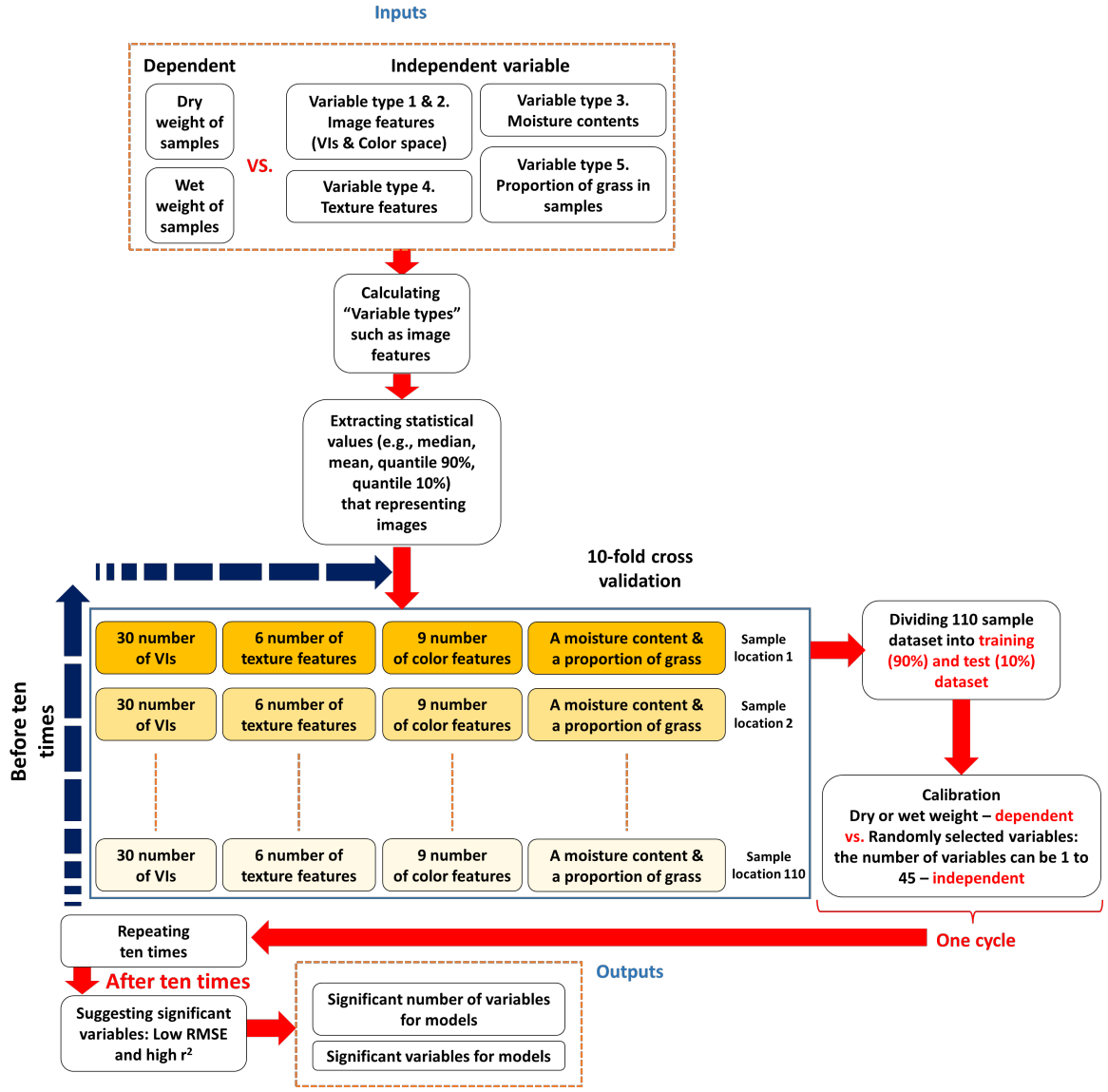

Fig 5. The RFE algorithm process, which recommends significant variables for each statistical group (e.g., median or mean).

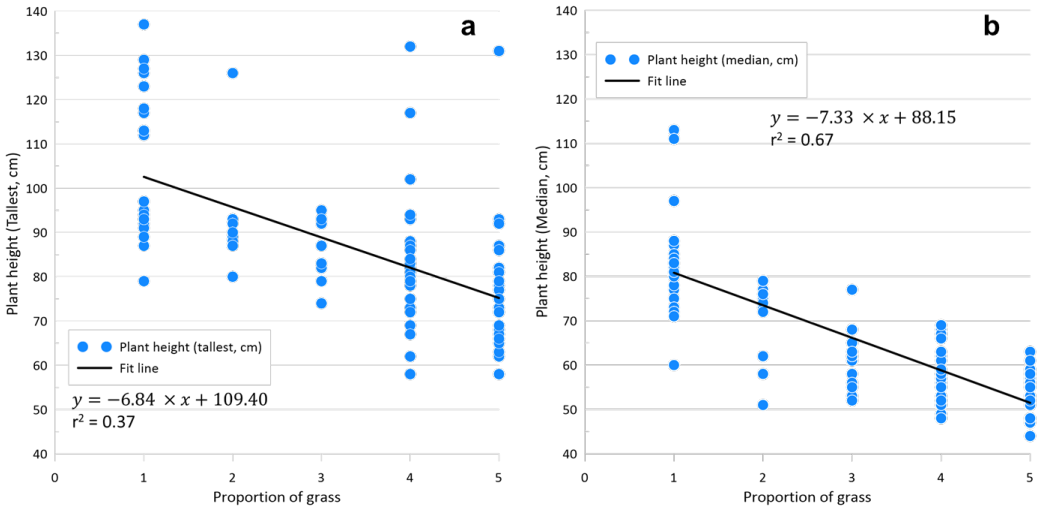

Fig 6. The relationship between proportion of grass in samples and tallest (a) and median (b) plant height in each quadrat.

RFE results

The RFE consisting of multiple linear and random forest regression models recommended a number of significant independent variables for training the machine learning models estimating either dry or wet mass based on the different summary statistics of the image data (Table 4). The linear regression RFE recommended fewer variables than the random forest type. On the other hand, the random forest RFE was more flexible than the linear type, so the average number of recommended variables for machine learning was four based on linear RFE and 36 based on random forest RFE types. Although the random forest RFE may offer a more accurate model, it can cause the model to be overfitted. Because the RFE algorithm considers not only correlation coefficient and RMSE, but also the coefficient of multiple correlation among independent variables, some model results that contain variables with weak correlation may be overfit.

Table 4. Suggested number of independent variables based on RFE results for different RFE algorithm types, dependent variables, and independent variable summarization methods.

Model type (RFE model type)	Linear		Random forest	
Dependent variable	Dry mass	Wet mass	Dry mass	Wet mass
Independent variable summarization method	Significant number of variables		Significant number of variables	
Maximum	1	1	47	45
Minimum	3	5	46	28
Mean	15	6	48	47
Median	2	3	5	11
Quantile 10%	15	4	18	48
Quantile 20%	1	4	45	47
Quantile 30%	1	4	47	39
Quantile 70%	1	11	46	48
Quantile 80%	1	10	49	48
Quantile 90%	1	1	11	6
Average	4.1	4.9	36.2	36.7

Based on the RFE analysis, we selected the “Mean” and “Quantile 90%” datasets from the random forest regression group for further examination. Pearson’s correlation coefficients (r) among the RFE-suggested independent variables are shown in Fig 7. The results showed that the direct relationship between each independent and dependent variables were not strong (i.e., $| r | < 0.4$), and higher correlations among independent variables were favorable for training machine learning models.

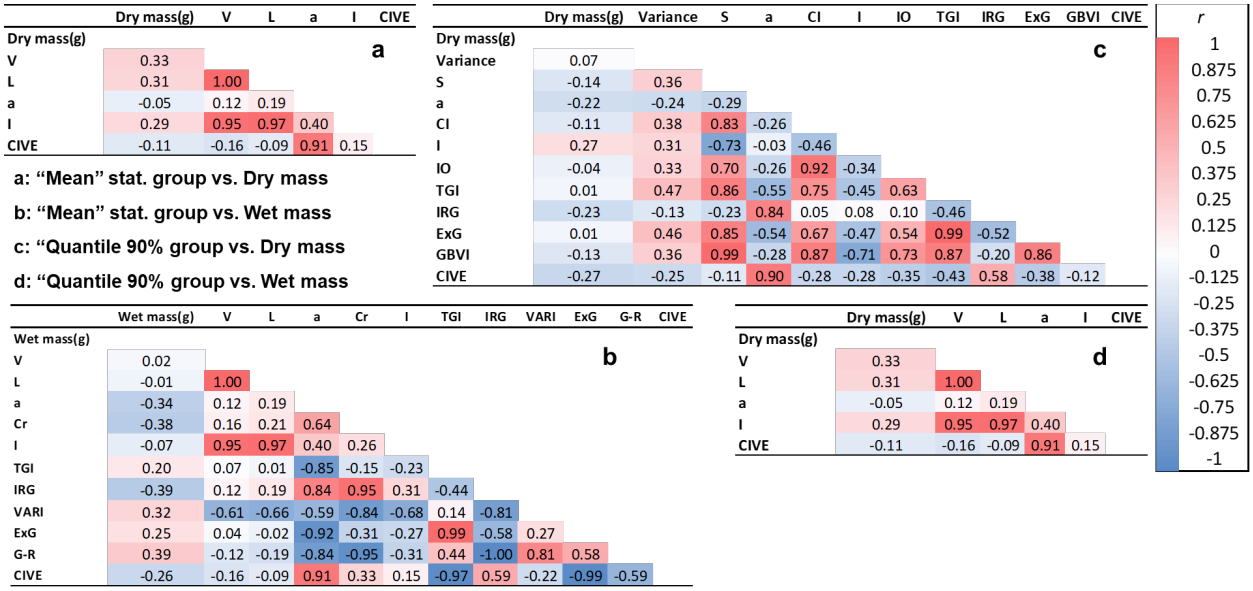

Fig 7. Heatmaps of Pearson's correlation coefficient (*r*) for selected RFE results: (a) Mean – dry with 5 independent variables (IV); (b) Mean – wet with 11 IV; (c) Quantile 90% – dry with 11 IV, and (d) Quantile 90% - wet with 6 IV (Variable abbreviations are defined in Table 2).

Random forest model test

Random forest regression was used to analyze the four candidate models with the independent variables suggested by the RFE (Table 5). The best-performing machine learning model (i.e., random forest regression) estimated wet mass based on the "Quantile 90%" statistical group of candidate variables from the random forest RFE (r^2 = 0.79, RMSE = 251.0 g/m^2, MAE = 187.6 g/m^2). Models predicting wet mass were more accurate than those for dry mass. This result might be due to the fact that the independent variables were extracted from images that were obtained in wet (fresh) condition. When comparing results from the two types of RFE (i.e., linear and random forest), models based on variables selected by the random forest RFE were more accurate. Therefore, additional analysis focused on the models using variables from the random forest RFE.

Table 5. Results of random forest regression on candidate independent variables provided by the two RFE types (linear and random forest).

	Dependent variables	Dry mass			Wet mass		
	Statistical group	r^2	RMSE (g/m^2)	MAE (g/m^2)	r^2	RMSE (g/m^2)	MAE (g/m^2)
Random forest RFE model type	Median	0.72	72.0	53.6	0.78	258.7	198.1
	Quantile 90%	0.75	71.6	51.9	0.79	251.1	187.6
Linear RFE model type	Median	0.59	90.5	66.9	0.66	312.9	233.9
	Quantile 10%	0.6	84.8	57.8	0.65	305.0	223.7

Scatter plots of predicted vs. measured quadrat mass for each model are shown in Fig 8. Points in the graphs are identified by the location (northern vs. roadside (western) area) and the proportion of the grass in the samples as described in Table 3. Noticeable outliers, single points detached from the fit line and the main group of points, generally were from the roadside area. According to Table 3, the main composition of most (36 out of 60) samples in the roadside area was red clover. These images were generally somewhat non-uniform, including shade, soil, or dried weeds, which may have caused the poor fit. On the contrary, the images in the northern area were generally more uniform which may have caused them to be better estimated by the model.

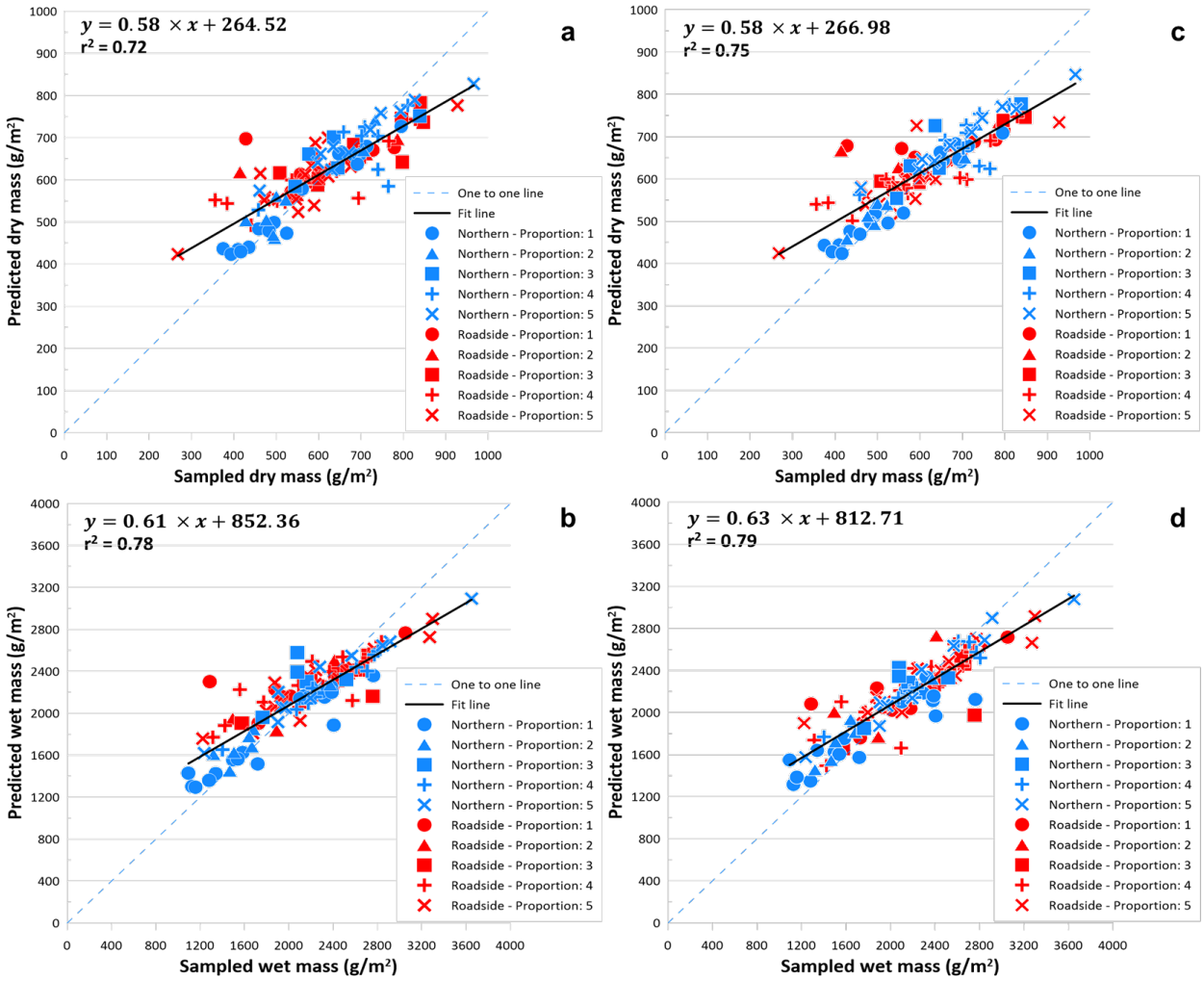

Fig 8. Results of random forest machine learning model. a: "Mean" group (independent variable) vs. dry mass (dependent variable) model, b: "Mean" group vs. wet mass, c: "Quantile 90%" group vs. dry mass, d: "Quantile 90%" group vs. wet mass. Definition of "Proportion" can be found in Table 3.

Discussion

The approach of using a single-frame imagery data acquisition to reduce data collection time and effort was quite successful (model r^2 = 0.72 to 0.79) although only RGB images and a simple random forest machine learning technique were used. Prior to configuring the random forest model and the RFE algorithm, the moisture content of the growing hay, the proportion of grass in the ground truth data samples, and image texture features were considered to be independent variables having high potential. Since a large portion of fresh (wet) hay mass is water, the proportion of grass is related to the wet (fresh) mass of the samples, and texture features help classify objects, we expected that these variables could be important. However, Table 6 shows that these three components (texture features, proportion of grass, moisture content of samples) contributed little to model training. While this finding is somewhat contradictory to findings of previous research, it does provide important information for guiding future data analysis efforts. Of course, it would be important to verify that the unimportance of these variables was not just due to the specific conditions of this study and could be replicated in other fields and data acquisitions.

Table 3. Frequency of independent variables used in the models for "Mean" dry and wet datasets and "Quantile 90%" dry and wet weight datasets.

	Independent variables that recommended by RFE				
	Color space	**Vegetation index**	**Texture features**	**Proportion of grass**	**Moisture content (for dry weight group)**
Frequency	10	22	1	0	0
The highest frequency variable	a (from L*a*b)	CIVE	Variance		

Outliers are further examined in Fig 9 for the "Quantile 90%" - dry mass model. Quadrat images are provided for several identified outliers, showing that the outlier quadrats were mainly composed of a mixture of grass and clover, or clover alone. These same quadrats were also identified as outliers in the other groups of models. Pearson's correlation coefficient (*r*) between the dry mass of samples and the proportion of the grass (i.e., group number in Table 3) in the roadside area, and the northern area were -0.10 and 0.48, respectively. This is the reason why the direct correlation (*r*) between dry or wet mass and independent variables was low (Fig 7). However, the high and robust correlation among independent variables (e.g., color features and vegetation indices) compensated for the low correlation to dependent variables. This implies that the uniform and evenly distributed northern area was more suitable for training the hay yield model.

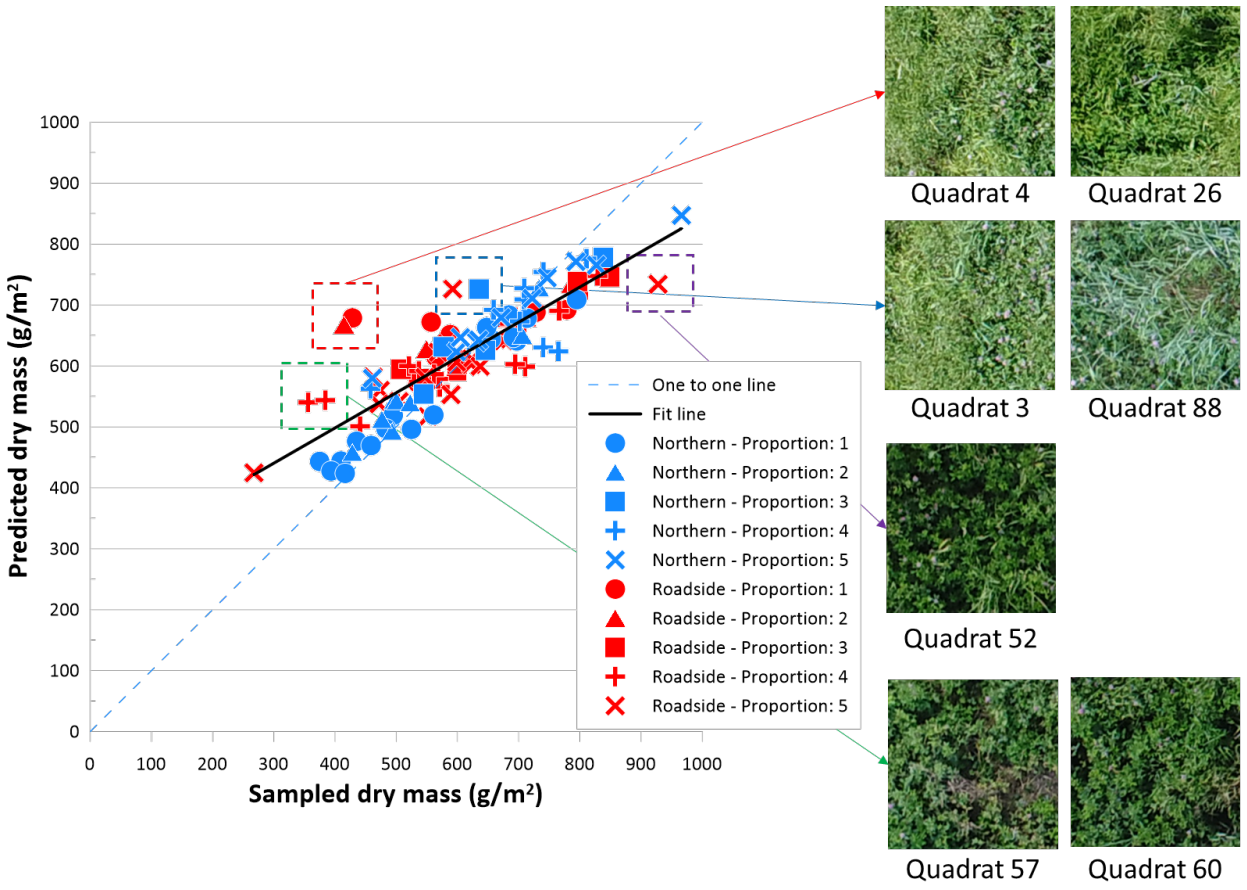

Fig 9. "Quantile 90%" - dry mass model results (left) and images of identified outliers (right). Definition of "Proportion" can be found in Table 3.

From the above results, the study showed that using a single image data collection and machine learning models based on RGB channels could predict hay yield while focusing on spatial accuracy and a rapid workflow. Moreover, this study provided information on which variables were more related to predicting wet (fresh) mass. In the future, applying multispectral imagery datasets in prediction models or using both RGB and multispectral imagery may lead to improved models. Also in future work, it will be important to focus on calibration sites with homogeneous vegetation to provide more accurate results.

Conclusion

Hay yield monitoring systems can provide farmers with spatial information in a field useful for site-specific management. As compared with previous methods used to estimate hay yield, improving the spatial resolution of yield information was the primary goal of this research.

This research investigated the use of remote sensing technologies for estimating hay yield. The specific objective was developing fast and spatially accurate models using a machine learning model (random forest) and testing suitable independent variables among five types of extracted variables from RGB images. The 110 calibration samples were collected in two areas (roadside and northern parts) within one day of harvest, immediately after imaging.

Candidate variables for modeling included measured vegetation moisture content, the proportion of grass in samples, vegetation indices, and texture features. From these, significant independent variables selected by the RFE algorithm and were used to train random forest machine learning models. Results showed the potential of this approach, with the best model achieving $r^2 = 0.79$, RMSE = 251.05 g/m^2, and MAE = 187.59 g/m^2.

Future work to improve hay yield estimation might include running deep learning models, collecting calibration data from locations with homogeneous vegetation characteristics, collecting enough samples to run machine learning models, and application of multispectral images.

References

Abdulridha, J., Batuman, O., & Ampatzidis, Y. (2019). UAV-based remote sensing technique to detect citrus canker disease utilizing hyperspectral imaging and machine learning. *Remote Sensing*, *11*(11), 1373. https://doi.org/10.3390/rs11111373

Castro, W., Marcato Junior, J., Polidoro, C., Osco, L. P., Gonçalves, W., Rodrigues, L., Santos, M., Jank, L., Barrios, S., Valle, C., Simeão, R., Carromeu, C., Silveira, E., Jorge, L. A. de C., & Matsubara, E. (2020). Deep learning applied to phenotyping of biomass in forages with UAV-based RGB imagery. *Sensors*, *20*(17), 4802. https://doi.org/10.3390/s20174802

de Oliveira, G. S., Marcato Junior, J., Polidoro, C., Osco, L. P., Siqueira, H., Rodrigues, L., Jank, L., Barrios, S., Valle, C., Simeão, R., Carromeu, C., Silveira, E., André de Castro Jorge, L., Gonçalves, W., Santos, M., & Matsubara, E. (2021). Convolutional neural networks to estimate dry matter yield in a guineagrass breeding program using UAV remote sensing. *Sensors*, *21*(12), 3971. https://doi.org/10.3390/s21123971

Feng, P., Wang, B., Liu, D. L., & Yu, Q. (2019). Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. *Agricultural Systems*, *173*, 303–316. https://doi.org/10.1016/j.agsy.2019.03.015

Ghamisi, P., Plaza, J., Chen, Y., Li, J., & Plaza, A. J. (2017). Advanced spectral classifiers for hyperspectral images: a review. *IEEE Geoscience and Remote Sensing Magazine*, *5*(1), 8–32. https://doi.org/10.1109/MGRS.2016.2616418

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-3*(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Huang, X., Zhang, L., Wang, B., Li, F., & Zhang, Z. (2018). Feature clustering based support vector machine recursive feature elimination for gene selection. *Applied Intelligence*, *48*(3), 594–607.

Maughan, J. D., Mathanker, S. K., Grift, T. E. & Hansen, A. C. (2012). Yield monitoring and mapping systems for hay and forage harvesting: a review. *2012 ASABE Annual International Meeting.* Dallas, Texas, July 29 - August 1, 2012. https://doi.org/10.13031/2013.41885

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Liu, P. (2015). A survey of remote-sensing big data. *Frontiers in Environmental Science*, *3*. https://www.frontiersin.org/article/10.3389/fenvs.2015.00045

Maguire, S., Godwin, R. J., Smith, D. F., & O'Dogherty, M. J. (2003). Hay and forage measurement for mapping. *Precision Agriculture: Papers from the 4th European Conference on Precision Agriculture, Berlin, Germany, 15-19 June 2003*, 379–384.

Osco, L. P., Ramos, A. P. M., Faita Pinheiro, M. M., Moriya, É. A. S., Imai, N. N., Estrabis, N., Ianczyk, F., Araújo, F. F. de, Liesenberg, V., Jorge, L. A. de C., Li, J., Ma, L., Gonçalves, W. N., Marcato Junior, J., & Eduardo Creste, J. (2020). A Machine learning framework to predict nutrient content in Valencia-orange leaf hyperspectral measurements. *Remote Sensing*, *12*(6), 906. https://doi.org/10.3390/rs12060906

Rueda-Ayala, V. P., Peña, J. M., Höglind, M., Bengochea-Guevara, J. M., & Andújar, D. (2019). Comparing UAV-Based Technologies and RGB-D Reconstruction methods for plant height and biomass monitoring on grass ley. *Sensors*, *19*(3), 535. https://doi.org/10.3390/s19030535