

WHEAT BIOMASS ESTIMATION FROM UAV IMAGERY USING AN ENSEMBLE LEARNING APPROACH WITH BAYESIAN OPTIMIZATION

F. Moradi¹, A. Zarei^{1*}, S. Ranjbar², S. Homayouni³

¹ Department of Geomatics, University College of Engineering, University of Tehran, Tehran, Iran – (fatemeh.moradii, arastou.zarei@ut.ac.ir)

² Department of Biological Systems Engineering, University of Wisconsin-Madison, Madison, WI, USA - sranjbar@wisc.edu

³ Center Eau Terre Environnement, Institut National de la Recherche Scientifique, Quebec City, QC, Canada - saeid.homayouni@inrs.ca

Commission VI, WG VI/3

KEY WORDS: Wheat Biomass, UAV Images, Ensemble Learning, Bayesian Optimization, Feature Importance.

ABSTRACT:

Wheat is one of the most important food supply and food security globally, especially in developing countries. Therefore, predicting the performance and determining the factors that affect the production of this product is very important. Biomass is one of the crop's most important biophysical parameters, and its correct estimation can help improve accurate monitoring of growth and crop performance forecasting. With the recent advances in remote sensing, access to aerial images taken by unmanned aerial vehicles (UAV) for monitoring crops has been provided. This study investigates the potential of visible UAV images and the resulting vegetation indices to estimate the dry biomass of two types of Brazilian wheat. For this purpose, the performance of three regression algorithms, including Random Forest (RF), eXtreme Gradient Boosting (XGB), and Gradient Boosting Machine (GBM), to estimate wheat biomass was evaluated. Also, to improve the performance of regression models, Bayesian optimization (BO) was used to adjust the Hyper-parameters, and random forest feature selection was used to select the optimal subset of features. Based on the results, the XGB algorithm with the Root Mean Square Error (RMSE) of about 911.86 (Kg/ha) and coefficient of determination (R^2) of about 0.89% showed better performance in biomass estimation than other algorithms.

1. INTRODUCTION

According to global research, the world's population will grow by 66% by 2050, and food security will become a significant challenge in agriculture due to this growth (Besthorn, 2013; Jalali et al., 2021). Wheat is crucial for food security; thus, timely monitoring of its growing status aids in maintaining sustainable agriculture (Ali et al., 2015a). Biomass is an essential indicator for crop evaluation of yield, grain quality, and gross primary production (Yue et al., 2017). Therefore, crop biomass estimation is essential for monitoring crop growth, improving crop management efficiency, and forecasting crop yield (Liu et al., 2019). Also, monitoring biomass production is a method for interpreting and evaluating the need for fertilizer, particularly nitrogen(N) deficiency in crops (Cilia et al., 2014).

Field measurements and remote sensing data are the two types of biomass estimation methodologies. Field measuring is an expensive and time-consuming method that can only be used in small-scale surveys (Du et al., 2019). Due to its high spatial resolution, consistency and cost-effectiveness, remote sensing technology is very effective in precision agriculture, especially for estimating crop biomass accurately (Dadras Javan et al., 2019; Ren & Feng, 2015). Recent advances in remote sensing have provided access to aerial images collected by unmanned aerial vehicles (UAV). The most important feature of UAVs is preparing aerial images with the very high spatial and temporal resolution desired by the researcher. In addition, the operational cost and complexity are much lower than other remote sensing platforms (Moradi et al., 2022). Also, compared to other UAV sensors, RGB sensors have been

considered due to their low cost in precision agriculture (Acorsi et al., 2019; Moradi et al., 2021).

Based on research conducted in the last two decades, biomass estimation and biophysical parameters of products from remote sensing images have been performed using statistical models (Yue et al., 2019). These models fall into two main categories: first, machine learning techniques (such as artificial neural networks (ANN), random forest regression (RF), and support vector machine (SVM)), and second technique: Conventional regression methods (such as multiple linear regression (MLR), stepwise multiple regression (MSR) and partial least squares regression (PLSR)) (Ranjbar, Akhoondzadeh, et al., 2021; Ranjbar, Zarei, et al., 2021). The most common method for estimating biomass using remote sensing data is vegetation indices and forming statistical models (Ali et al., 2015b). Kross et al. found a positive and significant correlation between maize, soybean biomass, and vegetation indices Green-NDVI, RVI, MTVI1, and NDVI from SPOT and Landsat images (Kross et al., 2015). Gao et al. Suggested that maize biomass could be estimated using the NDVI, RVI, and EVI vegetation indices from Chinese HJ-1A/B satellite imagery (Gao et al., 2013). Wang et al. Used HJ satellite images and 15 vegetation indices to estimate wheat biomass and compared the three algorithms ANN, SVR, and RF and showed that RF performed better than the other two algorithms (Wang et al., 2016). Gahrouei et al. used UAVSAR data to estimate three crops' biomass and leaf area index and compared ANN and MLR performance (Reisi-Gahrouei et al., 2019). Researchers also evaluated 2020 the potential for using vegetation indices derived from RapidEye multi-temporal data and two ANN and SVR techniques to estimate crop biomass

* Corresponding author

and leaf area index (Reisi Gahrouei et al., 2020). Pranga et al. used the features obtained from UAV visible and multispectral images and three RF, SVR, and PLSR algorithms to estimate the dry grassland biomass, which according to the results of the RF algorithm, had the best performance in biomass estimation (Pranga et al., 2021).

A review of previous studies shows that satellite remote sensing data is increasingly being used to monitor and estimate agricultural metrics. According to previous research, this study aims to use visible UAV images and ensemble learning algorithms to estimate wheat biomass with various genotypes and phenotypes. The main objectives of this research include 1) investigating the potential of UAV visible images to estimate wheat biomass and 2) assessing and comparing three ensemble machine learning regression techniques based on Bayesian optimization to forecast wheat biomass, including XGB, GBM, and RF.

2. CASE STUDY AND DATASET

2.1 Study Area

The study area is an experimental wheat field in southern Brazil with geographical coordinates (51° 40' W, 30° 6' S) consisting of several rectangular patches measuring 2.5 m by 1 m, which are plots with two types of Brazilian wheat (Schreiber et al., 2022). The genotypes used are TBIO Toruk and BRS Parrudo (48 Toruk plots and 40 Parrudo plots in Figure 1). Variation in crop growth was created for all test areas to receive a different amount of nitrogen. Nitrogen (N) rates have been used to diversify crop growth, assess biomass response, and grain yield to nitrogen availability called phenotypic diversity.

2.2 Data Collection

Data collection has been done in two stages; in the first stage, biomass was collected manually to obtain the ground truth data of the earth. The data were collected from May to October 2018. Dry biomass has been collected in three stages of wheat growth, including six fully expanded leaves called V6, three nodes, and a flowering stage in 0.27 square meters per plot. The collected plants were dried and weighed at 65 ° C, then the values per kg/h were extrapolated. In the second stage, the images were taken 50 meters above the ground using a sensor connected to the DJI Matrice 100 Quadcopter. The sensor used is DJI X3 Visible (RGB) with a resolution of 12 MB and a depth of 8-bit pixels. 80% frontal and 70% side overlap were considered for image collection with a resolution of 2.14 cm²/pix.

3. PROPOSED METHOD

In this research, the dry biomass of two varieties of Brazilian wheat was forecasted via UAV images utilizing the RF, GBM, and XGB ensemble learning algorithms (ELAs). According to

Figure 2, first, the spectral bands and vegetation indices obtained from RGB images were extracted, and then using the random forest feature selection algorithm, an optimal subset of these features was selected and considered as the input of ELAs. Input data were randomly divided into training and testing, which (70%) and (30%) data were considered for training and testing of ELAs, respectively. In order to increase the accuracy and performance of ELAs, the Bayesian optimization method was used to fine-tune the hyper-parameters of the three algorithms. Figure 2 shows the three processes involved in the implementation: 1) feature extraction, 2) biomass estimation, 3) accuracy assessment, and performance comparison of three ensemble learning algorithms in biomass prediction.

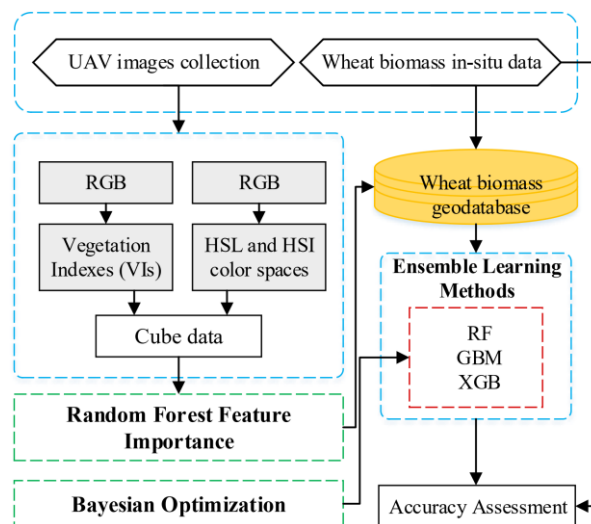


Figure 2. Flowchart of the proposed method.

3.1 Feature Extraction

Vegetation indices are a combination of two or more spectral bands effectively used to detect vegetation and have been used in extensive research to classify vegetation and monitor droughts and environmental changes (Yue et al., 2018). Also, many researchers have demonstrated that Vegetation indices can be used to estimate biophysical and biochemical parameters of crops, such as leaf area index (LAI), biomass, grain yield, and nitrogen accumulation (Lu et al., 2019). In this section, according to previous research (Lu et al., 2019; Maimaitijiang et al., 2019), the best Vegetation indices were used to predict biomass and were calculated for each plot (Table 1). Each plot's region of interest (ROI) was selected to remove pixels containing soil. For each ROI, the average digital number of RGB bands was also computed, as well as changing the RGB colour space to HSL (Hue, Saturation, Lightness) and HSI (Hue, Saturation, Intensity) (Table 2).

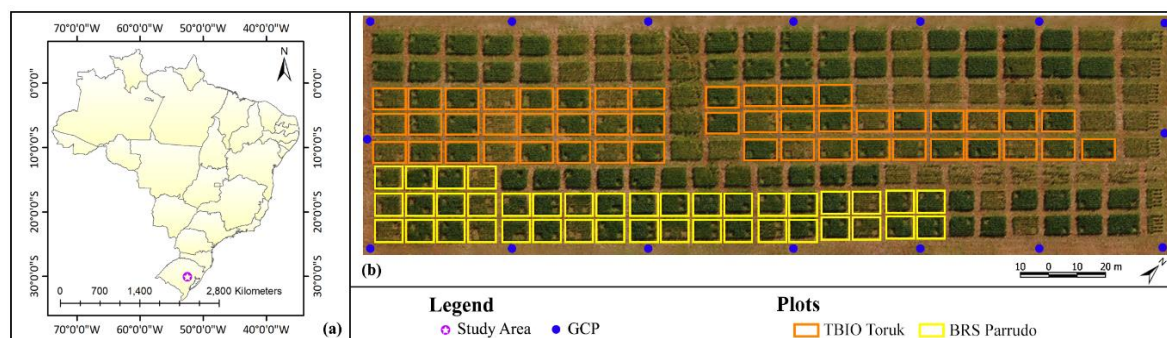


Figure 1. Schematic overview of (a) the study site location in Eldorado do Sul, Brazil, and (b) the experimental plot design with TBIO Toruk and BRS Parrudo varieties (<https://data.mendeley.com/datasets/3ntkg88d4d/1>).

Index	Formula
GLI	$(2 \times G - R - B)/(2 \times G + R + B)$
VARI	$(G - R)/(G + R - B)$
EXR	$1.4 \times r - g$
EXG	$2 \times g - r - b$
EXGR	$EXG - EXR$
VEG	$G/R^a B^{(1-a)}, a = 0.667$
CIVE	$0.441R - 0.881G + 0.385B + 18.78745$
COM1	$EXG + ExGR + CIVE + VEG$
COM2	$0.36EXG + 0.47CIVE + 0.17VEG$
RI	$R - G/R + G$
SCOM	$0.25EXG + 0.30ExGR + 0.33CIVE$
MEXG	$1.262G - 0.884R - 0.311B$
IF	$2 \times R - G - B/G - B$
TGI	$-0.5 [(\lambda_r - \lambda_b) (Rr - Rg) - (\lambda_r - \lambda_g) (Rr - Rb)]$
NGRDI	$(G - R)/(G + R)$

*R, G, and B represent the digital number of red, green, and blue channels. $r=R/(R+G+B)$, $g=G/(R+G+B)$, $b=B/(R+G+B)$

Table 1. Summary of vegetation indices used in this study.

Visible image spectral bands and color space	Feature	
	R	HSL-S
	G	HSL-L
	B	HSI-I
HSL-H		

Table 2. Features extracted from visible UAV images.

3.2 Regression analysis

3.2.1 Random Forest (RF)

RF is a powerful supervised ML method that was proposed by (Breiman, 1984) and has been widely used in RS and GIS applications, such as image classification (Stumpf & Kerle, 2011) and landslide susceptibility mapping (Ghorbanzadeh et al., 2019). This method is based on decision trees and operates by constructing many decision trees during the training process, making it less sensitive to over-fitting issues (Ghorbanzadeh et al., 2019). In the RF method, each decision tree generates outputs, and output weights derived from the votes are dedicated. The advantages of RF are that it is easy to apply because it requires only a few parameters, and it yields higher accuracy than other ML methods due to the bagging process (Rahmati et al., 2019). Additionally, it can deal with high-dimensional and complex data structures (Biau & Scornet, 2016). Hyper-parameters are very important for model optimization that involve 1) n estimator: the number of trees in the forest that will be optimized using the gride search method, 2) min samples leaf: The minimum number of samples required to be at a leaf node. A split point at any depth will only be considered if it leaves at least min leaf training samples in each left and right branch. This may have the effect of smoothing the model, especially in regression, and 3) max depth: this parameter determines how deep each tree grows in each reinforcement period.

3.2.2 Gradient Boosting Machine (GBM)

The GBM approach is used to transform weak learning trees into powerful ones. Each new tree in boosting is inserted into a modified version of the original dataset. To put it another way, trees are built sequentially in boosting so that each successive tree is reduced to the errors of the prior trees. Each tree learns from its predecessors and corrects any faults that remain. As a result, the tree that grows next learns from a previously updated version (Friedman, 2001). The primary distinction between the GBM and the random forest methods is that the random forest makes each tree independent. The GBM technique, on the other hand, accepts trees as an add-on (group) and introduces a weak tree to fix the weaknesses of existing susceptible trees in a step-by-step way, while the random forest approach merges the outcomes after the process. GBM, on the other hand, mixes the outcomes along the path. The gradient rise can outperform random forests if appropriately tuned parameters (Bahrami et al., 2021; Zarei et al., 2021).

3.2.3 eXtreme Gradient Boosting (XGB)

The XGB method is a type of GBM that searches for the optimal tree model using the most precise approximation. Gradient Boosting is the method's basic computation foundation, and it has two advantages over GBM: the first is execution speed, and the second is model performance. Increasing gradient trees with XGB is one of the quickest methods. The algorithm accomplishes this by exploiting the Gradient Boosting method's significant flaws by considering the cost function for all possible divisions when generating a new member. XGB overcomes this limitation by looking at the distribution of features across all data points on a sheet and using that knowledge to narrow the search space for prospective feature sharing. Calculating second-order gradients, or the cost function's second-order partial derivatives, gives you more information about the slopes' direction and how to get to the minimal cost function (Chen & Guestrin, 2016). In addition to the two parameters stated in the random forest technique (n estimator, max features, and max depth), two other essential parameters in XGB and GBM consist of 1) learning rate: this parameter defines the proportion of each tree in the final result and affects the pace of the algorithm in response to slope changes, and 2) max features: the number of features that should be considered when searching for the best subsets, Which is optimally defined in the training process as "auto", "sqrt", "log2" based on the number of input features.

3.2.4 Hyperparameter Tuning

Hyper-parameters must be initialized in any machine learning algorithm before creating a model. Precise adjustment of model Hyper-parameters maximizes model performance and accuracy (Yang & Shami, 2020). Manual tuning of Hyper-parameters is made by trial and error, which is very time-consuming. Another method for selecting Hyper-parameters is the use of optimization methods, the most common of which are Grid Search (GS), random search (RS), and Bayesian optimization (BO) (Arabi et al., 2022). The BO method is more accurate and faster than the GS and RS methods because it detects Hyper-parameters in each iteration by analyzing the values in the previous iteration (Hutter et al., 2019). BO is an advanced method used to tune the Hyper-parameters of the deep learning network. It has recently been used to adjust the Hyper-parameters of machine learning (Hutter et al., 2019), so the BO method has been used in this research. The optimization process in BO is made up of four key

components: search space, objective function, surrogate model, and acquisition function. Instead of splitting the data into two independent training and validation sets, the BO approach uses K-fold cross-validation to evaluate and update hyper-parameters using a validation dataset. The training data is divided into k parts in the k-fold. One of the k parts is considered evaluation data, and the remaining part is considered training data during k distinct phases. Then the average values of evaluation results are determined.

3.3 Accuracy Assessment

In this research, the R^2 and RMSE has been used to validate and compare the results of three regression models (Equations 1 and 2).

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

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

where N = the total number of observations
 y_i = vector of observed values
 \bar{y} = mean of the observed variables
 \hat{y}_i = vector of predicted dependent variables

4. RESULT AND DISCUSSION

This section analyzes the potential of an optimal subset of features derived from visible UAV images and the performance of three regression models for estimating wheat biomass. Also, the results of fine-tuning of hyper-parameters of the three regression models are given by BO, each of which is discussed separately in the following sections.

4.1 Features Selection

According to Tables (1, 2), 22 features were extracted from UAV images to estimate wheat biomass, which in Figure 3 shows the Pearson coefficient correlation matrix, the relationship between the extracted features and biomass. Some of the extracted features such as COM2, VEG, MEXG,

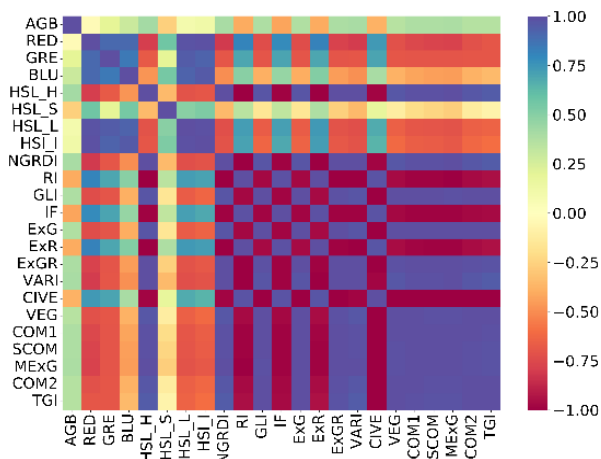


Figure 3. Pearson's correlation coefficient (r) between wheat biomass and features derived from UAV visible images.

SCOM, and COM1, in addition to AGB, have a high correlation with each other, and if all the features are used, there will be a problem of multicollinearity and data redundancy. It seriously affects regression performance and runtime, so selecting the optimal subset of the extracted features is necessary. The random forest feature selection algorithm is a way to automatically select the optimal features that were used according to Figure 4, twelve features of higher importance including (Blue and green bands, changing the color space of HSL-L, HSL-S, HSL-I, HIS-I, and, vegetation indices of TGI, EXR, IF, VARI, NGRDI, RI).

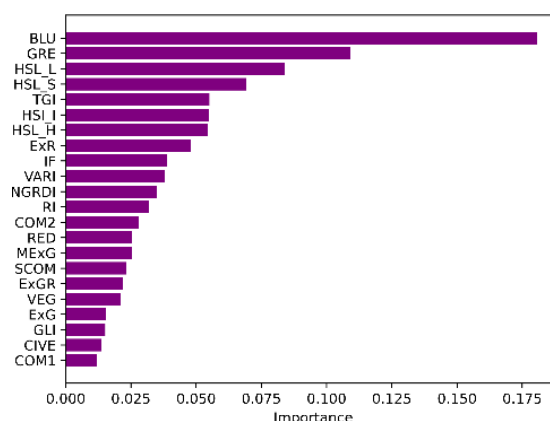


Figure 4. Random forest for feature importance on a regression problem.

4.2 Bayesian Optimization

This study used cross-validation and BO to fine-tune the hyper-parameters. In cross-validation, the training sets were divided into ten folds. Each time, one of the folds was used for validation, and the remaining nine-fold was used for training. Also, according to Figure 5, 150 iterations were considered to adjust the hyper-parameters.

The hyper-parameters adjust for GBM, and XGB includes: Shrinks the contribution of each tree (learning rate), The number of boosting stages to conduct (n_estimators), Limits the number of nodes in the tree (max_depth), and the number of features to consider when searching for the best split (max_features). Also, four hyper-parameters n_estimators, max_depth, max_features, and the minimum number of samples required to be in a leaf nod (min_samples_leaf) were adjusted for the random forest regression algorithm. Figure 5 shows the BO results and optimal values for each hyper-parameter with an asterisk; the horizontal axis represents the number of iterations, and the vertical axis represents the search space for the hyperparameter values. Also, the value of max_features is considered in three regression algorithms equal to the root of the features.

4.3 Accuracy of the Reconstructed Models

To compare the performance of three regression machine learning algorithms for estimating wheat biomass, validation values are given in Table 3. The XGB algorithm with RMSE = 911.86 and $R^2 = 0.89$ has the best performance compared to other algorithms. In fact, according to a recent study (Zarei et al., 2021), for two major causes, the XGB algorithm performs better than the other two algorithms: First, in the XGB algorithm, by calculating second-order gradients, the loss function's second partial derivatives provide more information on the gradient direction and how to get to the loss

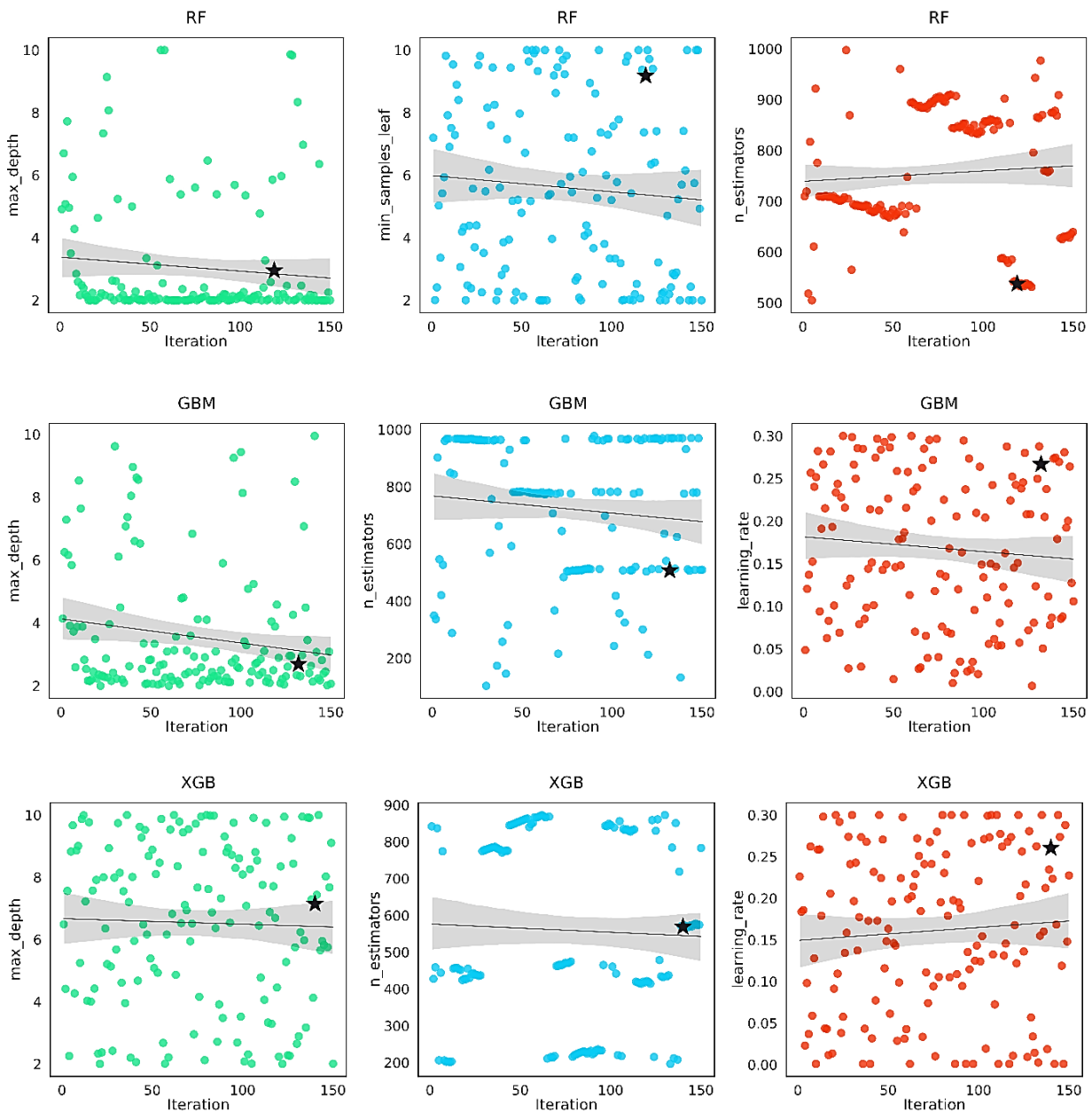


Figure 5. Results of fine-tuning hyperparameters using Bayesian optimization.

function's minimum. Furthermore, advanced regularization (L1, L2) improves model generalization. According to Table 3, although the RMSE and R2 values of XGB and GBM are almost similar, the learning of XGB is much faster than GBM. Similarly, the performance of XGB in terms of execution time is better than the other two algorithms.

ELAs-Regression	R ² (%)	RSME (kg/h)	Runtime (min)
RF	0.78	1279.59	55
GBM	0.87	994.24	18
XGB	0.89	911.86	5

Table 3. Accuracy assessment of three ELAs in estimating dry AGB.

According to scatter plots in Figure 6, the closer the distribution of points to the line ($x = y$), the higher the accuracy of wheat biomass estimation. The vertical axis of the scatter plots shows the predicted biomass, which is 500 to 1000 (kg/h) for the V6 stage, 800 to 4000 (kg/h) for the three-node stage, and 4000 to 10000 (kg/h) for the flowering stage. According to Figure 6, all three regression models accurately estimated the biomass in the initial stage of wheat growth. However, these steps are less accurately predicted because the vegetation indices are saturated in the medium to high canopy cover in the second and third stages of wheat growth and wheat biomass. Nevertheless, the accuracy of the proposed method for estimating wheat biomass is higher than or equal to the accuracy of previous research methods. Wang et al. used HJ satellite images and 15 vegetation indices to estimate wheat biomass and compared the performance of three ANN, SVR, and RF algorithms to estimate wheat biomass in three growth stages and showed

that the RF algorithm in three stages of Wheat growth, respectively with RMSE 477, 1126.2 and 1808.2 has the best performance compared to other algorithms (Wang et al., 2016). In our research, although the RF algorithm performed worse than the other two regression algorithms compared to previous research, for two main reasons was able to predict wheat biomass with higher accuracy: 1) Since the images obtained from the UAV have a higher spatial resolution; as a result, biomass estimation is done with higher accuracy, and 2) Due to the fine-tune of hyper-parameters based on BO, the performance of regression algorithms has improved compared to previous research.

In another study, researchers used visible images from UAV to estimate the biomass of two types of Brazilian wheat (TBIO Toruk and BRS Parrudo) and compared the

performance of the ANN and Convolution neural network (CNN) algorithms. Based on the results of this study, ANN and CNN algorithms could estimate wheat biomass with RMSE values of 826.4, 940.5, and R^2 values of 0.9056% and 0.9065%, respectively (Schreiber et al., 2022).

5. CONCLUSION

Due to the direct relationship between biomass and crop yield, biomass can be used indirectly to estimate wheat yield and predict overproduction or shortage of wheat. In this study, to estimate the dry biomass of wheat with different phenotypes and genotypes, the visible images of the UAV and the resulting vegetation indices, as well as three machine learning algorithms (RF, GB, XGB), have been used. Finally, the main results obtained in this research are as follows. 1) UAV Imagery to estimate biomass is a low-cost and fast data collection method that provides more information and accuracy. 2) Among machine learning techniques, the XGB algorithm with $RMSE = 911.86$ and $R^2 = 0.89$ has the best performance and accuracy in biomass estimation. It is also faster than the other two algorithms. 3) The fine-tuning of the hyperparameters improves the performance of machine learning algorithms. 4) Comparing the accuracy of the XGB algorithm with other studies that used deep learning methods to estimate biomass showed that our results are close to the results of other researchers and can be promising.

To sum up, the proposed method is simple and fast. However, despite the high accuracy, deep learning methods require many training data and expensive GPUs, which requires a lot of time and money. Despite the efficiency and high speed of the proposed method in estimating wheat biomass, in the second and third stages of wheat growth, the forecast was less accurate, so it is suggested that in future research, in addition to vegetation indices, other features should be used.

ACKNOWLEDGMENT

The authors would like to express their deep gratitude to Prof. Schreiber, Atkinson Amorim, and Parraga from the Rio Grande do Sul State University for providing the Brazilian wheat dataset.

REFERENCES

- Acorsi, M. G., das Dores Abati Miranda, F., Martello, M., Smaniotto, D. A., & Sartor, L. R. (2019). Estimating Biomass of Black Oat Using UAV-Based RGB Imaging. *Agronomy*, 9(7), 344. <https://doi.org/10.3390/agronomy9070344>
- Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., & Notarnicola, C. (2015a). Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sensing*, 7(12), 16398–16421. <https://doi.org/10.3390/rs71215841>
- Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., & Notarnicola, C. (2015b). Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sensing*, 7(12), 16398–16421. <https://doi.org/10.3390/rs71215841>
- Arabi, S., Eshtehardian, E., & Shafiei, I. (2022). Using Bayesian Networks for Selecting Risk-Response Strategies in Construction Projects. *Journal of Construction*

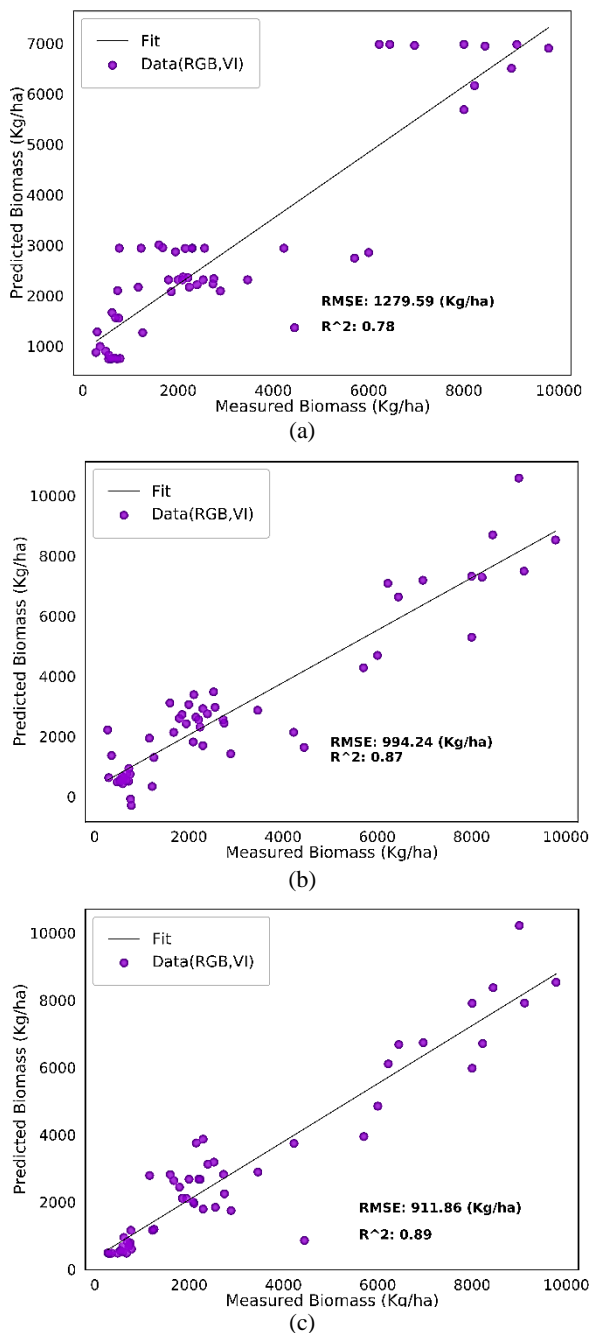


Figure 6. Scatterplots of wheat biomass estimates with three ELAs: (a) Random Forest, (b) Gradient Boosting Machine, and (c) eXtreme Gradient Boosting

Engineering and Management, 148(8), 04022067.
[https://doi.org/10.1061/\(asce\)co.1943-7862.0002310](https://doi.org/10.1061/(asce)co.1943-7862.0002310)

Bahrami, H., Homayouni, S., Mahdianpari, M., & Safari, A. (2021). ENSEMBLE LEARNING FOR CROP MONITORING FROM MULTITEMPORAL OPTICAL AND SYNTHETIC APERTURE RADAR EARTH OBSERVATIONS. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 5929–5932. <https://doi.org/10.1109/IGARSS47720.2021.9555121>

Besthorn, F. H. (2013). Vertical Farming: Social Work and Sustainable Urban Agriculture in an Age of Global Food Crises. *Australian Social Work*, 66(2), 187–203. <https://doi.org/10.1080/0312407X.2012.716448>

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>

Breiman, Leo. (1984). *Classification and regression trees*. Wadsworth International Group.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>

Cilia, C., Panigada, C., Rossini, M., Meroni, M., Busetto, L., Amaducci, S., Boschetti, M., Picchi, V., & Colombo, R. (2014). Nitrogen Status Assessment for Variable Rate Fertilization in Maize through Hyperspectral Imagery. *Remote Sensing*, 6(7), 6549–6565. <https://doi.org/10.3390/rs6076549>

Dadras Javan, F., Mortazavi, S., Moradi, F., & Toosi, A. (2019). NEW HYBRID PAN-SHARPENING METHOD BASED ON TYPE-1 FUZZY-DWT STRATEGY. *Pdfs.Semanticscholar.Org*. <https://doi.org/10.5194/isprs-archives-XLII-4-W18-247-2019>

Du, Y., Wang, J., Liu, Z., & Lin, Y. (2019). Estimation and multiscale transformation of aboveground biomass: An hgsu-oriented approach based on multisource data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7), 2388–2396. <https://doi.org/10.1109/JSTARS.2019.2921817>

Friedman, J. H. (2001). 999 REITZ LECTURE GREEDY FUNCTION APPROXIMATION: A GRADIENT BOOSTING MACHINE 1. In *The Annals of Statistics* (Vol. 29, Issue 5). <https://www.jstor.org/stable/2699986>

Gao, S., Niu, Z., Huang, N., & Hou, X. (2013). Estimating the Leaf Area Index, height and biomass of maize using HJ-1 and RADARSAT-2. *International Journal of Applied Earth Observation and Geoinformation*, 24(1), 1–8. <https://doi.org/10.1016/j.jag.2013.02.002>

Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S., Tiede, D., & Aryal, J. (2019). Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection. *Remote Sensing*, 11(2), 196. <https://doi.org/10.3390/rs11020196>

Hutter, F., Kotthoff, L., & Vanschoren, J. (Eds.). (2019). *Automated Machine Learning*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-05318-5>

Jalali, J., Ahmadi, A., & Abbaspour, K. (2021). Runoff responses to human activities and climate change in an arid watershed of central Iran. *Hydrological Sciences Journal*, 66(16), 2280–2297. <https://doi.org/10.1080/02626667.2021.1985724>

Kross, A., McNairn, H., Lapen, D., Sunohara, M., & Champagne, C. (2015). Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *International Journal of Applied Earth Observation and Geoinformation*, 34(1), 235–248. <https://doi.org/10.1016/j.jag.2014.08.002>

Liu, Y., Liu, S., Li, J., Guo, X., Wang, S., & Lu, J. (2019). Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Computers and Electronics in Agriculture*, 166, 105026. <https://doi.org/10.1016/j.compag.2019.105026>

Lu, N., Zhou, J., Han, Z., Li, D., Cao, Q., Yao, X., Tian, Y., Zhu, Y., Cao, W., & Cheng, T. (2019). Improved estimation of aboveground biomass in wheat from RGB imagery and point cloud data acquired with a low-cost unmanned aerial vehicle system. *Plant Methods*, 15(1), 17. <https://doi.org/10.1186/s13007-019-0402-3>

Maimaitijiang, M., Sagan, V., Sidike, P., Maimaitiyiming, M., Hartling, S., Peterson, K. T., Maw, M. J. W., Shakoor, N., Mockler, T., & Fritsch, F. B. (2019). Vegetation Index Weighted Canopy Volume Model (CVM VI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 151, 27–41. <https://doi.org/10.1016/j.isprsjprs.2019.03.003>

Moradi, F., Javan, F. D., & Samadzadegan, F. (2022). Potential evaluation of visible-thermal UAV image fusion for individual tree detection based on convolutional neural network. *International Journal of Applied Earth Observation and Geoinformation*, 113, 103011. <https://doi.org/10.1016/j.jag.2022.103011>

Moradi, F., Javan, F., & Toosi, A. (2021). *Tree detection using UAV based imagery system based on Random Forest classification*. <https://sciforum.net/paper/view/10819>

Pranga, J., Borra-Serrano, I., Aper, J., de Swaef, 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 Sensing*, 13(17), 3459. <https://doi.org/10.3390/rs13173459>

Rahmati, O., Choubin, B., Fathabadi, A., Coulon, F., Soltani, E., Shahabi, H., Mollaefar, E., Tiefenbacher, J., Cipullo, S., Ahmad, B. bin, & Tien Bui, D. (2019). Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods. *Science of the Total Environment*, 688, 855–866. <https://doi.org/10.1016/j.scitotenv.2019.06.320>

- Ranjbar, S., Akhondzadeh, M., Brisco, B., Amani, M., & Hosseini, M. (2021). Soil Moisture Change Monitoring from C and L-band SAR Interferometric Phase Observations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 7179–7197. <https://doi.org/10.1109/JSTARS.2021.3096063>
- Ranjbar, S., Zarei, A., Hasanlou, M., Akhondzadeh, M., Amini, J., & Amani, M. (2021). Machine learning inversion approach for soil parameters estimation over vegetated agricultural areas using a combination of water cloud model and calibrated integral equation model. *Journal of Applied Remote Sensing*, *15*(01), 018503. <https://doi.org/10.1117/1.jrs.15.018503>
- Reisi Gahrouei, O., McNairn, H., Hosseini, M., & Homayouni, S. (2020). Estimation of Crop Biomass and Leaf Area Index from Multitemporal and Multispectral Imagery Using Machine Learning Approaches. *Canadian Journal of Remote Sensing*, *46*(1), 84–99. <https://doi.org/10.1080/07038992.2020.1740584>
- Reisi-Gahrouei, O., Homayouni, S., McNairn, H., Hosseini, M., & Safari, A. (2019). Crop biomass estimation using multi regression analysis and neural networks from multitemporal L-band polarimetric synthetic aperture radar data. *International Journal of Remote Sensing*, *40*(17), 6822–6840. <https://doi.org/10.1080/01431161.2019.1594436>
- Ren, H., & Feng, G. (2015). Are soil-adjusted vegetation indices better than soil-unadjusted vegetation indices for above-ground green biomass estimation in arid and semi-arid grasslands? *Grass and Forage Science*, *70*(4), 611–619. <https://doi.org/10.1111/gfs.12152>
- Schreiber, L. V., Atkinson Amorim, J. G., Guimarães, L., Motta Matos, D., Maciel da Costa, C., & Parraga, A. (2022). Above-ground Biomass Wheat Estimation: Deep Learning with UAV-based RGB Images. *Applied Artificial Intelligence*, *36*(1). <https://doi.org/10.1080/08839514.2022.2055392>
- Stumpf, A., & Kerle, N. (2011). Object-oriented mapping of landslides using Random Forests. *Remote Sensing of Environment*, *115*(10), 2564–2577. <https://doi.org/10.1016/j.rse.2011.05.013>
- 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 Journal*, *4*(3), 212–219. <https://doi.org/10.1016/j.cj.2016.01.008>
- Yang, L., & Shami, A. (2020). On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice. In *Elsevier*. <https://www.sciencedirect.com/science/article/pii/S0925231220311693>
- 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 Sensing*, *10*(2), 66. <https://doi.org/10.3390/rs10010066>
- Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., Feng, H., & Xu, B. (2017). Estimation of Winter Wheat Above-Ground Biomass Using Unmanned Aerial Vehicle-Based Snapshot Hyperspectral Sensor and Crop Height Improved Models. *Remote Sensing*, *9*(7), 708. <https://doi.org/10.3390/rs9070708>
- Yue, J., Yang, G., Tian, Q., Feng, H., Xu, K., & Zhou, C. (2019). Estimate of winter-wheat above-ground biomass based on UAV ultrahigh-ground-resolution image textures and vegetation indices. *ISPRS Journal of Photogrammetry and Remote Sensing*, *150*, 226–244. <https://doi.org/10.1016/j.isprsjprs.2019.02.022>
- Zarei, A., Hasanlou, M., & Mahdianpari, M. (2021). A comparison of machine learning models for soil salinity estimation using multi-spectral earth observation data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *5*(3), 257–263. <https://doi.org/10.5194/isprs-annals-V-3-2021-257-2021>