NDVI Prediction with RGB UAV Imagery Utilizing Advanced Machine Learning Regression Models

Ilyas Aydin [1], Umut Gunes Sefercik [1]

[1] GTU, Engineering Faculty, Dept. of Geomatics Engineering 41400, Kocaeli, Turkey - ilyasaydin@gtu.edu.tr, sefercik@gtu.edu.tr

Keywords: NDVI prediction, Multispectral UAV, RGB UAV, CatBoost, LightGBM, Stacking Ensemble model.

Abstract

The ever-evolving technology has significantly affected the sensors used in UAV cameras and has played an important role in the expansion of the application areas of hobbyist and commercial UAVs. In particular, UAVs with multispectral (MS) cameras, which have the potential to detect a wide range of spectral information, are widely used in many popular research areas such as precision agriculture and forestry. However, despite their advanced capabilities, the high cost of these technologies limits their accessibility for basic users. In this study, the agricultural potential of RGB UAVs, which have a much wider user base due to their lower cost, was investigated by predicting the Normalized Difference Vegetation Index (NDVI), which is widely preferred for plant classification, growth and health monitoring. In the literature, RGB camera-based NDVI prediction studies involving machine learning and deep learning algorithms have focused on the correlation of the results with the reference data (R^2) or the model accuracy of the algorithms used. The approaches applied have generally been tested on single photographs or solely on vegetation areas. In this study, using the MS UAV NDVI map as reference, a comprehensive evaluation approach was applied where each pixel of the NDVI prediction maps produced by categorical boosting (CatBoost), light gradient boosting machine (LightGBM) and a stacking ensemble learning model obtained from the combination of both algorithms, whose performance in NDVI estimation has not been tested extensively before. The models were tested in an urban area with numerous buildings and a large study area with dense vegetation. The performance of the NDVI maps was analyzed using R^2, Root Mean Square Error (RMSE), Normalized Median Absolute Deviation (NMAD) and Standard Deviation (STD) metrics. As a result of the comprehensive analysis, it was found that the models performed similarly in general, but the LightGBM model was slightly behind the others. The considerable results around 0.81-0.83 as R^2 and ~0.09 as RMS and STD clearly showed that RGB cameras can be a lower-cost alternative solution for generating NDVI maps in agricultural studies when supported by machine learning models.

1. Introduction

Today, unmanned aerial vehicles (UAVs) have diverse applications, including disaster monitoring, land cover and land use (LULC) mapping, crop health assessment, urban heat map production, and various hobby uses (Ahmed et al., 2008; Do et al., 2018; Cho et al., 2023; Saponaro and Tarantino, 2022). Their potential to record high-resolution data from lower altitudes has paved the way for their use, especially for monitoring agricultural fields. Agricultural practices are gaining importance day by day in order to increase productivity. This application area, which forms the basis for many research topics, makes frequent use of images recorded by advanced cameras and indices derived from them. The normalized difference vegetation index (NDVI), which provides information on vegetation and plant health, is among the main resources commonly used in precision agriculture practices (Houborg and McCabe, 2016; Mahajan and Bundel, 2016). This index, calculated as the ratio of the near-infrared and red bands, cannot be obtained with standard camera equipment that detects in the visible spectrum range (0.4 μm - 0.7 μm). This situation brings forth the need for multispectral (MS) cameras, which increases the cost.

MS UAVs have proven effectiveness in detecting diseases and monitoring crop development, thanks to the high-quality data they provide. However, their high cost makes them less accessible to many users. Due to this reason, UAVs with relatively cheaper RGB cameras have started to be tested for their suitability for agricultural use. Initially, the generation of RGB-based indices was attempted and such as the triangular greenness index (TGI) were developed (Raymond Hunt et al., 2011). However, vegetation reflectance is not high in the visible wavelength range and the difference between bands does not provide reliable data, making this approach ineffective. Fuentes-Peailillo et al. reported that the TGI index produced inaccurate results compared to the NDVI index in determining vegetation areas in their study based on the comparison of RGB-based indices (Fuentes-Peailillo et al., 2018).

To date, only a limited number of studies have been published on NDVI estimation using RGB imagery. Most of these studies focus exclusively on evaluating prediction performance within vegetation-covered areas (Moscovini et al., 2024; Wang et al., 2020). Most previous studies were based on point-based analyses of single images (Costa et al., 2020; Moscovini et al., 2024). In contrast, this study applied a model-based approach that allows the evaluation of each pixel on an NDVI map generated from RGB data, and an NDVI map generated from an MS UAV was used as a reference in the analyses.

In this study, we tested the performance of light gradient boosting machine (LightGBM) and categorical boosting (CatBoost) (Latha and Bommi, 2023; Xiang et al., 2022), which are tree-based ensemble learning models widely used to solve regression problems, on NDVI estimation using the RGB UAV dataset. In addition to these models, an ensemble stacking model was used as a third approach, integrating the estimation results of both LightGBM and CatBoost. As a result of the analysis, it has been revealed that the applied approaches can produce considerable results in NDVI estimation from RGB images. Thus, it has been determined that low-cost RGB UAV data can serve as an alternative source in the production of NDVI maps required in agricultural research.

2. Study Area and Materials

The study area is located in the northern part of the Gebze Technical University campus in Kocaeli, Türkiye (Figure 1). The campus area close to sea provides ideal conditions for this research, with dense vegetation and human-made features. Since the study area has many land cover classes, it contains a wide range of NDVI values belonging to these classes.

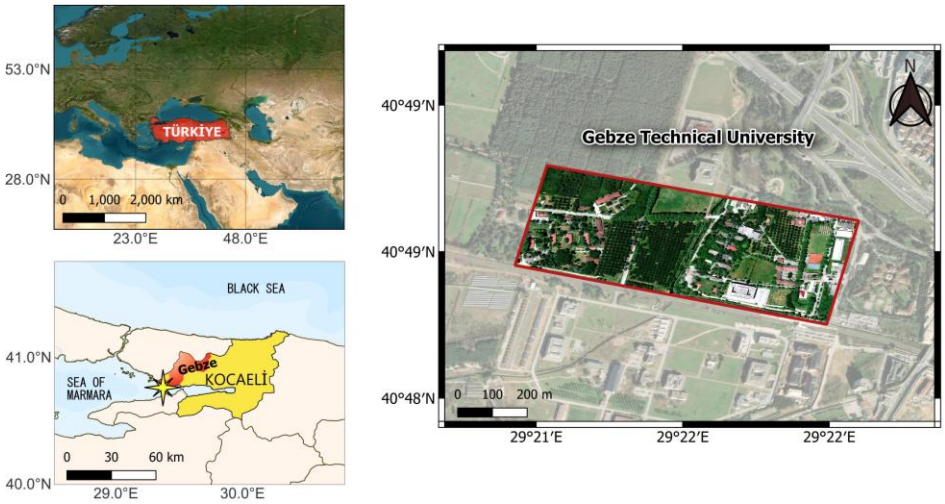


Figure 1. Study area.

Aerial images were acquired with DJI Phantom 4 MS UAV and their radiometry was calibrated by MAPIR V2 calibration target. The 5.74 mm focal length UAV camera used in the study consists of six different sensors as five single bands (R, G, B, NIR, RED) and one composite (RGB) sensor. CMOS sensors with a spectral range of 450-840 nm and a spatial resolution of 2.08 MP can capture six different images simultaneously for each photo taken.

3. Methodology

The methodology for producing NDVI prediction map directly from RGB UAV images begins with land reconnaissance, followed by the image acquisition. In the preparation of the data to be used for regression, two separate flights were conducted in order to obtain the train and test datasets.

The flights were planned in DJI Ground Station Pro software, which offers autonomous flight support. Both flights were conducted in polygonal geometry by utilizing same parameters as an altitude of 110 m and 80% front and 60% side overlap ratios. The nadir view (90°) was preferred to ensure maximum radiometric accuracy, and the aerial photos were achieved with 0.06 m ground sampling distance (GSD).

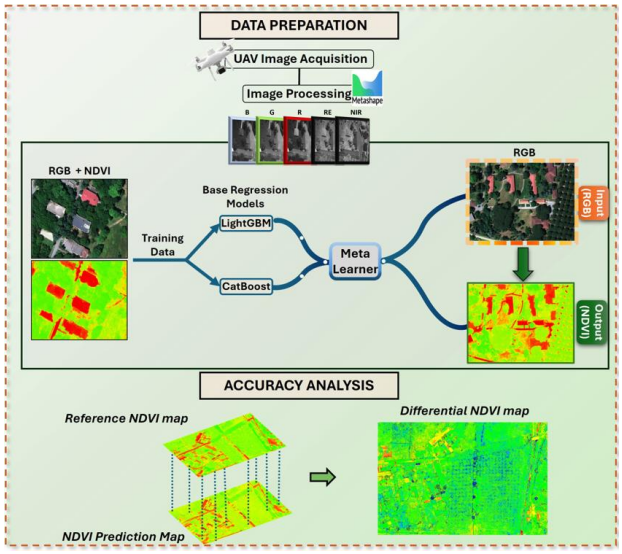


Figure 2. The followed workflow for data preparation and accuracy analysis.

In photogrammetric processing of the UAV imagery, Agisoft Metashape Pro, a Structure from Motion (SfM)-based software, was utilized. To prepare the test and training datasets, identical parameters were used in geometric and radiometric corrections, dense point cloud generation and digital surface model (DSM) production. Image matching and dense point cloud generation were performed using the high-quality option. While the image captured by the UAV equipped with RTK hardware was automatically matched based on SfM with high precision, eight mobile 0.25 m x 1 m polycarbonate ground control points (GCPs) installed in a radial distribution across the study area, were used for geometric accuracy control. As a result, the entire photogrammetric workflow was conducted in a controlled manner, resulting in outputs with sub - pixel accuracy.

First, a MS orthomosaic of the entire study area was produced by stacking all mono-bands obtained from two separate flights. Then, the train and test regions were masked on the orthomosaic and NDVI maps of the relevant areas were produced. In this way, reference maps were obtained to be used for testing model performances and accuracy of the final products. Then, all images from the composite RGB camera were processed and the train and test datasets were generated by extracting the areas within the mask boundaries from the 3-band orthomosaic (Figure 3). The boundary of the training dataset was carefully defined to ensure it included examples of all LULC classes. Additionally, the impact of increasing data size on training time, which varies depending on hardware specifications, was mitigated by identifying the most suitable area for the training dataset (~20.5 million pix.).

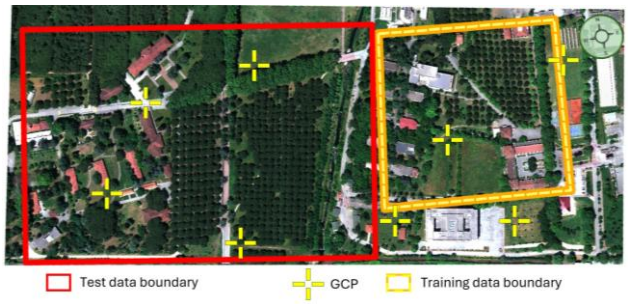


Figure 3. Training - Test data boundary and GCP distribution over the study area.

After preparing the training data, the decision tree-based LightGBM and CatBoost algorithms, which have not been

widely used in similar studies, were employed as base models for NDVI estimation.

The LightGBM algorithm is an important algorithm that is frequently used in solving regression problems as it speeds up the training process designed with large data sets (Shehadeh et al., 2021; Xuan et al., 2023). The leaf-wise tree growth strategy in LightGBM enables a rapid reduction in error values during training, enhancing model performance. However, this approach often results in deeper tree structures, which can increase the risk of overfitting to the training data. This problem is avoided by adjusting the maximum tree depth and maximum tree leaves parameters provided to the user (Zhu et al., 2022).

CatBoost regression is widely regarded for its high performance in various forecasting and prediction tasks (Qian et al., 2021; Zhong et al., 2023). This algorithm begins by constructing a weak learner and iteratively trains additional learners using the outcomes of previous rounds. The primary goal is to enhance the learner's accuracy by reducing bias during the training process. Finally, all the weak learners from each iteration are weighted and combined to form a strong and accurate model.

NDVI values were computed using an stacking ensemble learning model that integrates predictions from both the CatBoost and LightGBM models. Stacking ensemble learning is a powerful machine learning algorithm that improves the overall prediction accuracy by evaluating the prediction results of various base models (Shu et al., 2022). The fundamental principle of a stacking model is to combine the outputs of base models and use them as input features for a meta-model, which is trained to enhance the overall prediction performance. In this approach, for each pixel in the input RGB data, the NDVI values predicted by CatBoost and LightGBM are utilized as training inputs for the meta-model. To enhance the accuracy of the final predictions, a ridge regressor is employed as the final predictor in the meta-model. Ridge regression is widely used in statistical modelling and is highly effective in handling highly correlated independent variables (Jha et al., 2024). It is frequently preferred for evaluating and weighting multiple predictions in stack models (Khooran et al., 2023).

In machine learning regression algorithms, hyperparameter optimization is essential for revealing true model performance. It is a powerful approach for identifying the optimal values of hyperparameters, ensuring the algorithm performs at its best. In this study, the Optuna algorithm was used to estimate the hyperparameters in the models. This algorithm is a promising alternative due to its success in hyperparameter optimization for tree-based models (Lai et al., 2023).

When applying the Optuna algorithm, the first step is to define the parameters to be estimated in the CatBoost model and establish the general structure of the model. The optimization process proceeds iteratively, with Optuna suggesting different sets of hyperparameters and their corresponding value ranges. The model is trained using these suggested parameters, and its performance is evaluated using the RMSE metric on the test data. A total of 50 different parameter sets were tested, and trials that performed poorly during training were stopped early through pruning. Finally, the best-performing hyperparameter values were identified, along with an output showing the impact of each parameter on the model's performance. The same workflow was applied for the LightGBM algorithm. Figure 4 shows the hyperparameter importance determined by Optuna for both CatBoost and LightGBM.

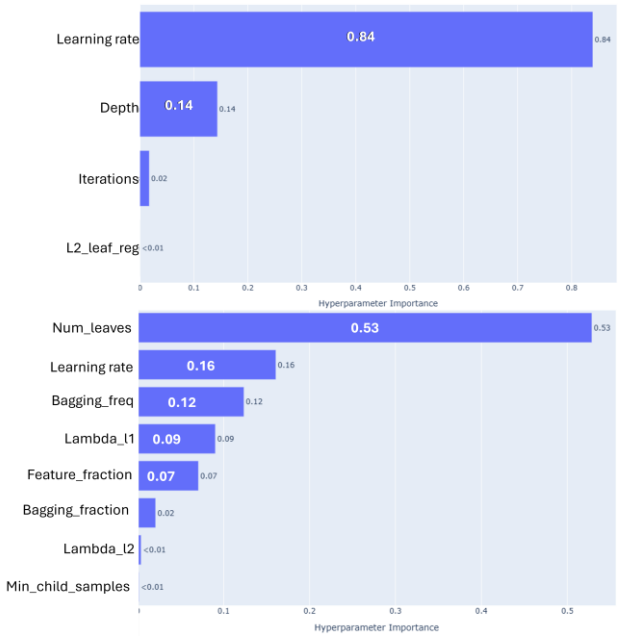


Figure 4. Hyperparameter importance in CatBoost (upper) and LightGBM (bottom) regression models.

3.1 Accuracy Analysis

The accuracy analysis was conducted in two stages: first, by evaluating the performance of the regression models, and second, by assessing the quality of the final outputs. In evaluating the model performance, the training data was divided into 90% train and 10% test, and the RMS values calculated for the test data set were used as performance measures.

Prior to the accuracy assessment, the planimetric alignment between the reference data and the production results was evaluated, confirming that sub-pixel accuracy was achieved. NDVI prediction map produced by regression models were evaluated using accuracy metrics such as Root Mean Squared Error (RMSE), Standard Deviation (STD), Normalized Median Absolute Deviation (NMAD), and the coefficient of determination (R^2), as shown in Equations 1-4. For visual analysis and interpretation, differential NDVI maps were generated to represent the differences between the reference data and the NDVI prediction maps, using the formulation provided in Equation 5. When calculating the standard deviation, the differences in NDVI values ($\Delta NDVI_i$) between the reference data and the model results are used, along with the mean of these differences (μ). n represents the total number of pixels considered in the calculation. R^2 and RMSE are calculated based on the comparison between the actual values (y_i), the predicted values ($\hat{y}_i$), and the mean of the actual values ($\bar{y}$). The NMAD metric, which is less sensitive to outliers than the standard deviation, provides robust results about the overall distribution of the data.

$$STD_{NDVI} = \sqrt{\frac{\sum_{i=1}^{n}(\Delta NDVI_i - \mu)^2}{n-1}} \quad (1)$$

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

$$NMAD = 1.4826 \times median(|\, X_i - median(X)\,|) \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (4)$$

$$\text{NDVI}_{\text{Differential}} = \text{NDVI}_{\text{Reference}} - \text{NVDI}_{\text{Regression}} \quad (5)$$

4. Results

The fact that the train and test data were obtained from flights performed for independent areas increased the interpretability in the generalization of the results. Figure 5 shows the orthomosaic and reference NDVI map produced with MS UAV data for the test area (~19 ha) to be applied for the performance analysis of the regression models.

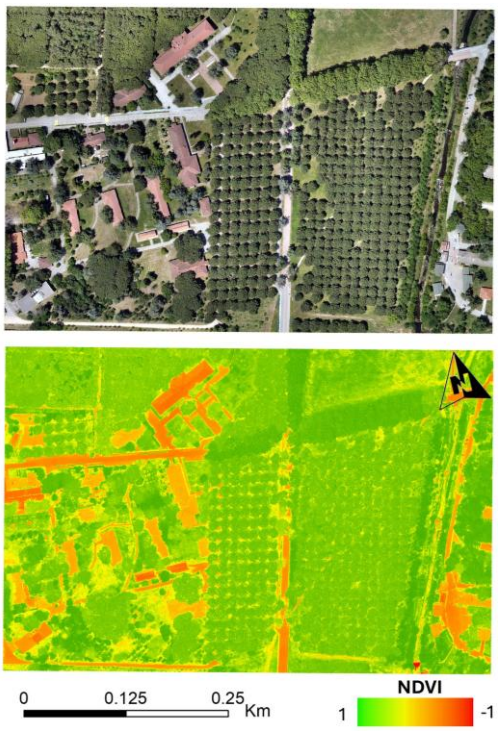
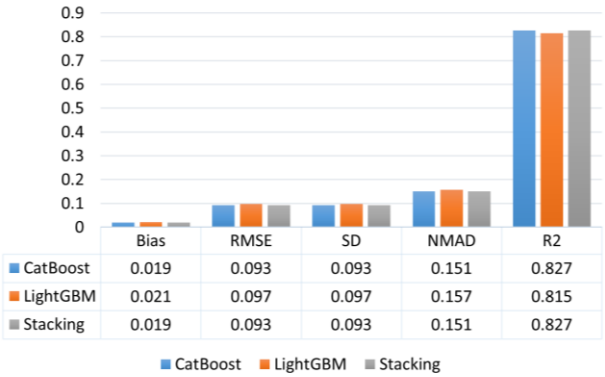


Figure 5. Orthomosaic (upper) and reference MS NDVI map (bottom) for the test area.

Figure 6 shows the performance of Catboost, LightGBM and stacking regression models on the test data, as determined by accuracy metrics. All models exhibited very similar bias values, with only minor differences. The LightGBM model exhibited slightly higher systematic errors than the other models; however, the difference was minimal. Bias values identified prior to the calculation of other metrics were eliminated, ensuring consistency and enhancing the interpretability of the results. The stacking ensemble model and CatBoost demonstrate the lowest RMSE values (0.093), indicating better prediction accuracy compared to LightGBM (0.097). CatBoost and the stacking model achieve the highest R^2 values (0.827), indicating better model fit and higher explained variance. LightGBM lags slightly behind with an R^2 of 0.815. Statistical evaluations revealed that, as a general result, CatBoost achieved higher accuracy compared to LightGBM in model comparisons. The stacking ensemble model performed as well as CatBoost, with no significant performance differences.



	Bias	RMSE	SD	NMAD	R2
CatBoost	0.019	0.093	0.093	0.151	0.827
LightGBM	0.021	0.097	0.097	0.157	0.815
Stacking	0.019	0.093	0.093	0.151	0.827

Figure 6. Model-based accuracy assessment of regression results.

Figure 7 shows the NDVI prediction maps obtained by applying CatBoost, LightGBM and stacking models. While all models were able to clearly distinguish between vegetation and other class, the predicted NDVI values for the classes showed small differences between the regression models.

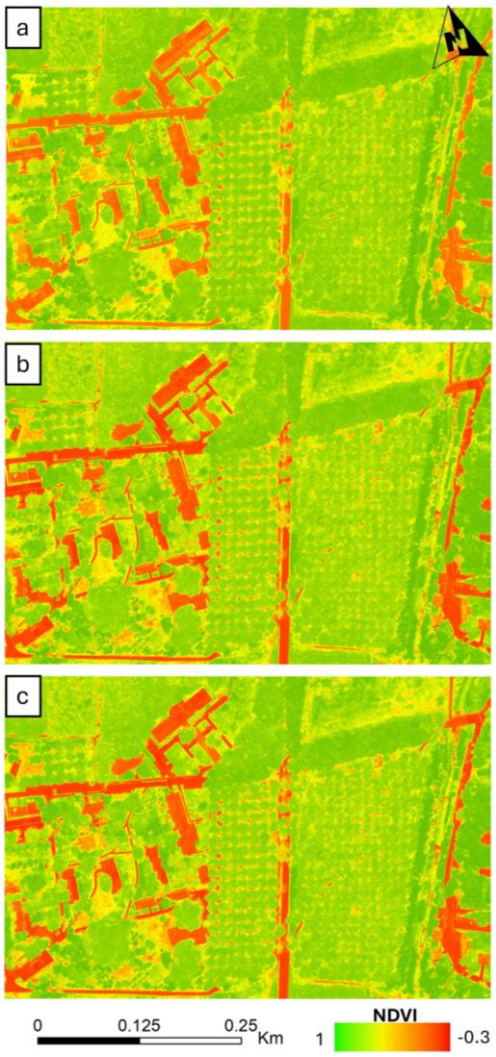


Figure 7. NDVI regression results: (a) LightGBM, (b) CatBoost, (c) Stacking model.

Figure 8 illustrates the differential maps, highlighting the pixel-to-pixel differences between the reference data and the NDVI

maps generated by the regression models. When the differential maps are examined, it is clearly seen that the entire study area can be represented by an NDVI difference of ± 0.06, expressed in green. Given the emphasis on analyzing vegetation areas and the objective of distinguishing these areas from other land cover classes, the results were deemed satisfactory. Specifically, the NDVI values corresponding to vegetation areas were predicted with a minimal error rate (± 0.06) across all models. Conversely, the model performance declined in the building, soil, and road classes, which are characterized by low NDVI values. Additionally, the LightGBM model demonstrated lower performance compared to other models. In Figure 8, sections of the soil and building classes with incorrect predictions are highlighted by black frames to facilitate a visual comparison between the LightGBM (a) and CatBoost (b) models. In the building structure located in the north of the study region, the LightGBM model showed better results with small differences, while it was behind CatBoost in other regions and overall performance. The stack ensemble model showed largely the same performance as CatBoost as in the statistical results.

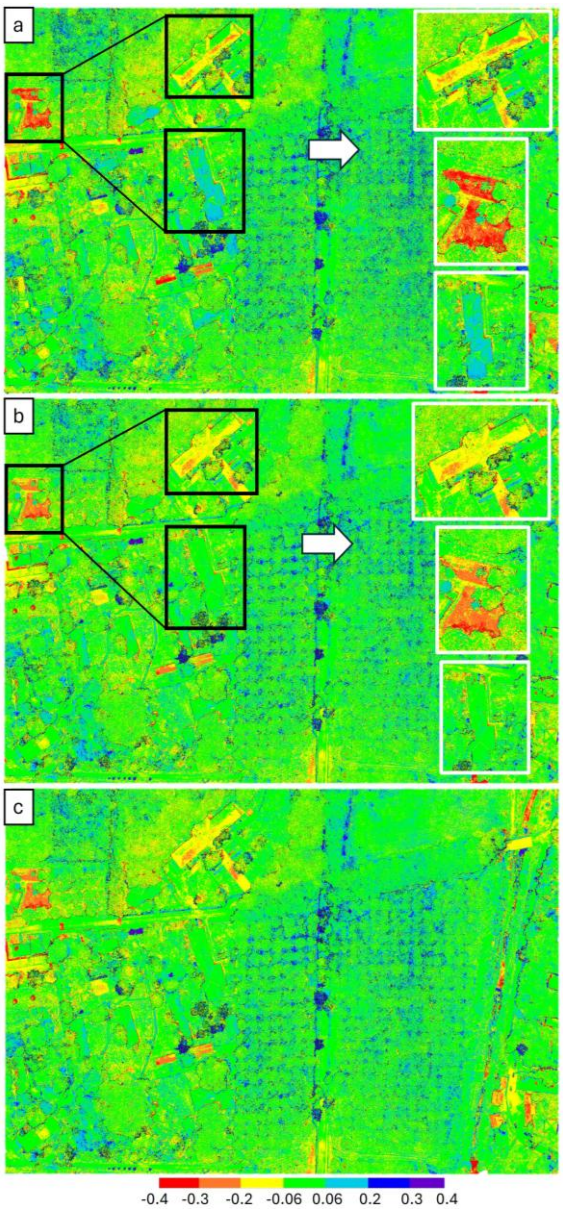


Figure 8. NDVI differential map results: (a) LightGBM, (b) CatBoost, (c) Stacking ensemble model.

An analysis of the visual results reveals that regression models tend to have a higher likelihood of error when predicting classes characterized by low NDVI values, such as buildings and soil. In contrast, fewer errors were observed in classes associated with higher NDVI values. These findings suggest that the regression models face challenges in predicting low NDVI values and tend to overestimate the NDVI values for these classes, predicting them as higher than the actual values.

5. Conclusions

In this study, where current ML algorithms were evaluated for NDVI prediction, regression was performed with three different model approaches and all prediction maps were evaluated on a model basis. Although statistical evaluations revealed minimal differences with no significant results, visual assessments indicated significant disparities between the CatBoost and LightGBM algorithms. Therefore, statistical evaluation of the regression results alone proves insufficient, highlighting the importance of visual analysis.

The findings reveal that the stacking ensemble model does not yield a significant performance improvement and closely aligns with the prediction results of the CatBoost model. At this point, it has been observed that the stacking process highlights the high-performing algorithm and prioritizes its prediction results.

In the light of the analyses, it has been determined that the application of continuously evolving algorithms to sources with limited spectral information, such as RGB images, can yield promising results for agricultural studies. Furthermore, it has been demonstrated that the approaches employed could serve as a viable alternative for basic user groups, given their limited access to high-cost MS cameras.

References

Ahmed, A., Nagai, M., Tianen, C., Shibasaki, R., 2008. UAV based monitoring system and object detection technique development for a disaster area. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*., XXXVII, 37, 373-377.

Cho, Y. Il, Yoon, D., Lee, M.J., 2023. Comparative Analysis of Urban Heat Island Cooling Strategies According to Spatial and Temporal Conditions Using Unmanned Aerial Vehicles (UAV) Observation. *Applied Sciences*, 13(18), 10052. doi.org/10.3390/app131810052.

Costa, L., Nunes, L., Ampatzidis, Y., 2020. A new visible band index (vNDVI) for estimating NDVI values on RGB images utilizing genetic algorithms. *Computers and Electronics in Agriculture*, 172. doi.org/10.1016/j.compag.2020.105334.

Do, D., Pham, F., Bhandari, S., Raheja, A., 2018. Machine learning techniques for the assessment of citrus plant health using UAV-based digital images. *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping III*, 10664, 189-200. SPIE. doi.org/10.1117/12.2303989.

Fuentes-Peailillo, F., Ortega-Farias, S., Rivera, M., Bardeen, M., Moreno, M., 2018. Comparison of vegetation indices acquired from RGB and Multispectral sensors placed on UAV. *IEEE ICA-ACCA 2018 - IEEE International Conference on Automation/23rd Congress of the Chilean Association of Automatic Control: Towards an Industry 4.0 - Proceedings*. doi.org/10.1109/ICA-ACCA.2018.8609861.

Houborg, R., McCabe, M.F., 2016. High-Resolution NDVI from planet’s constellation of earth observing nano-satellites: A new data source for precision agriculture. *Remote Sens (Basel)*, 8. doi.org/10.3390/rs8090768.

Jha, A., Goel, V., Kumar, M., Kumar, G., Gupta, R., Jha, S.K., 2024. An Efficient and Interpretable Stacked Model for Wind Speed Estimation Based on Ensemble Learning Algorithms. *Energy Technology*, 12(6), 2301188. doi.org/10.1002/ente.202301188.

Khooran, M., Golbahar Haghighi, M.R., Malekzadeh, P., 2023. Remaining Useful Life Prediction by Stacking Multiple Windows Networks with a Ridge Regression. *Iranian Journal of Science and Technology, Transactions of Mechanical Engineering*, 47(2), 583-594. doi.org/10.1007/s40997-022-00526-9.

Lai, J.P., Lin, Y.L., Lin, H.C., Shih, C.Y., Wang, Y.P., Pai, P.F., 2023. Tree-Based Machine Learning Models with Optuna in Predicting Impedance Values for Circuit Analysis. *Micromachines*, 14(2), 265. doi.org/10.3390/mi14020265.

Latha, R., Bommi, R.M., 2023. Hybrid CatBoost Regression model based Intrusion Detection System in IoT-Enabled Networks. *Proceedings of the 9th International Conference on Electrical Energy Systems*, ICEES. doi.org/10.1109/ICEES57979.2023.10110148.

Mahajan, U., Bundel, B.R., 2016. Drones for Normalized Difference Vegetation Index (NDVI), to Estimate Crop Health for Precision Agriculture: A Cheaper Alternative for Spatial Satellite Sensors. *International Conference on Innovative Research in Agriculture, Food Science, Forestry, Horticulture, Aquaculture, Animal Sciences, Biodiversity, Ecological Sciences and Climate Change (AFHABEC-2016)*.

Moscovini, L., Ortenzi, L., Pallottino, F., Figorilli, S., Violino, S., Pane, C., Capparella, V., Vasta, S., Costa, C., 2024. An open-source machine-learning application for predicting pixel-to-pixel NDVI regression from RGB calibrated images. *Computers and Electronics in Agriculture*, 216, 108536. doi.org/10.1016/j.compag.2023.108536.

Qian, Q., Jia, X., Lin, H., Zhang, R., 2021. Seasonal forecast of nonmonsoonal winter precipitation over the eurasian continent using machine-learning models. *Journal of Climate*, 34(17), 7113-7129. doi.org/10.1175/JCLI-D-21-0113.1.

Hunt Jr, E.R., Daughtry, C.S.T., Eitel, J.U., Long, D.S., 2011. Remote sensing leaf chlorophyll content using a visible band index. *Agronomy Journal*, 103(4), 1090-1099. doi.org/10.2134/agronj2010.0395.

Saponaro, M., Tarantino, E., 2022. LULC Classification Performance of Supervised and Unsupervised Algorithms on UAV-Orthomosaics. *In International Conference on Computational Science and Its Applications*, 311-326. doi.org/10.1007/978-3-031-10545-6_22.

Shehadeh, A., Alshboul, O., Al Mamlook, R.E., Hamedat, O., 2021. Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, LightGBM, and XGBoost regression. *Automation in Construction*, 129, 103827. doi.org/10.1016/j.autcon.2021.103827.

Shu, M., Fei, S., Zhang, B., Yang, X., Guo, Y., Li, B., Ma, Y., 2022. Application of UAV Multisensor Data and Ensemble Approach for High-Throughput Estimation of Maize Phenotyping Traits. *Plant Phenomics*. doi.org/10.34133/2022/9802585.

Wang, L., Duan, Y., Zhang, L., Rehman, T.U., Ma, D., Jin, J., 2020. Precise estimation of NDVI with a simple NIR sensitive RGB camera and machine learning methods for corn plants. *Sensors*, 20(11), 3208. doi.org/10.3390/s20113208.

Xiang, W., Xu, P., Fang, J., Zhao, Q., Gu, Z., Zhang, Q., 2022. Multi-dimensional data-based medium- and long-term power-load forecasting using double-layer CatBoost. *Energy Reports*, 8, 8511-8522. doi.org/10.1016/j.egyr.2022.06.063.

Xuan, L., Lin, Z., Liang, J., Huang, X., Li, Z., Zhang, X., Zou, X., Shi, J., 2023. Prediction of resilience and cohesion of deep-fried tofu by ultrasonic detection and LightGBM regression. *Food Control*, 154, 110009. doi.org/10.1016/j.foodcont.2023.110009.

Zhong, W., Zhang, D., Sun, Y., Wang, Q., 2023. A CatBoost-Based Model for the Intensity Detection of Tropical Cyclones over the Western North Pacific Based on Satellite Cloud Images. *Remote Sensing*, 15(14), 3510. doi.org/10.3390/rs15143510.

Zhu, J., Su, Y., Liu, Z., Liu, B., Sun, Y., Gao, W., Fu, Y., 2022. Real-time biomechanical modelling of the liver using LightGBM model. *International Journal of Medical Robotics and Computer Assisted Surgery*, 18(6). doi.org/10.1002/rcs.2433.