The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-6-2025
ISPRS, EARSeL & DGPF Joint Istanbul Workshop "Topographic Mapping from Space" dedicated to Dr. Karsten Jacobsen's 80th Birthday
29–31 January 2025, Istanbul, Türkiye

A Comparative Analysis of Black-Box and Glass-Box Models for Poplar Plantation Mapping with Remote Sensing Data

Muhammed Yusuf Ozturk [1], Ismail Colkesen [1]

[1] Dept. of Geomatic Engineering, Gebze Technical University, Gebze – Kocaeli, Turkey – m.ozturk2020@gtu.edu.tr, icolkesen@gtu.edu.tr

Keywords: Poplar, XGBoost, Explainable Boosting Machine, SHAP, Glass-box, Feature Selection.

Abstract

Poplar trees are essential for industrial afforestation applications due to their globally recognized plantation practices, reputation, ability to produce a large quantity of raw material in a short time, diverse applications in wood production, suitability for hybridisation and breeding implementations, and the availability of various species and clones adapted to the soil and climate conditions. Accurate identification and mapping of poplar afforestation areas are therefore crucial for planners and decision-makers to manage inventory records and maximize economic value. In this study, the poplar tree mapping and feature selection performances of the glass-box Explainable Boosting Machine (EBM) algorithm were investigated using satellite images having different resolutions (i.e., Sentinel-2 and PlanetScope) and texture features. A robust black-box algorithm, XGBoost, was utilized as a benchmark algorithm to compare the performance of EBM. The results showed that the EBM algorithm outperformed the standard XGBoost algorithm by up to 2% in classifying poplar trees when both spectral bands and calculated texture features were used, for both satellite images. Additionally, using the high spatial resolution PlanetScope imagery resulted in a significant decrease in the classification accuracy of popular areas (about %10) compared to Sentinel-2 imagery. The study also assessed the most important features influencing the classification process. For this, while 15 features were selected employing the visual charts provided by EBM for interpreting the decision-making process, the SHAP technique was applied to examine the most prominent features in the XGBoost model structure. In this scenario, EBM and XGBoost presented greater performances for both satellite data compared. These findings emphasize EBM's consistent superiority, indicating that its enhanced interpretability can facilitate more precise feature selection and model refinement, particularly for Sentinel-2 imagery.

1. Introduction

The increasing demand for wood, coupled with restricted forest areas and protective measures, has highlighted the importance of fast-growing tree plantations as a solution to the challenges of sustaining wood supply and ecosystem security (Khayatinejad Aqgonbad et al., 2022; Colkesen et al., 2023). In particular, the advancement of industrial forestry and the application of afforestation with fast-growing tree species, especially poplar, have gained attention to meet the rising need for timber (Kocer and Kara, 2021). Poplars are favoured for their fast growth, easy production and hybridization, adaptability to diverse soils and climates, and extensive range of species, colonies, and varieties (Kavzoglu et al., 2024). Based on the National Poplar Commission Country Progress Report, encompassing the period between 2012 and 2015, with approximately 145,000 hectares dedicated to poplar plantations, Turkey is one of the leading countries in this sector (Velioglu and Akgul, 2016). Therefore, accurately detecting and mapping poplar plantations are essential for decision-makers and planners to maintain inventory records and optimize the economic and ecological benefits of poplar.

Obtaining valuable information about the object of interest using remote sensing technology and its primary output, satellite imagery, has emerged as a trending topic in large-scale forestry applications, especially tree species mapping and identifications. In this regard, substantial research has been conducted on extracting tree species by classifying satellite images. The selection of a suitable classification algorithm is one of the key issues affecting classification accuracy, among other factors such as the characteristics of the study area and the specifications of the satellite image. Recent advancements in machine learning have opened new opportunities for research in image processing and object detection, such as accurately identifying tree species and mapping their distribution. Extreme Gradient Boosting (XGBoost), one of the most popular decision tree-based algorithms developed by Chen and Guestrin (2016), has attracted considerable interest in classifying remotely sensed data due to its robustness. The main idea behind the XGBoost is to iteratively construct decision trees and correct the errors of the previous model in the current iteration. XGBoost has been widely used as a robust algorithm for its superior performance in many classification problems, but it is often categorized as a black-box model because its decision rules used in the prediction phase are unclear and easily interpretable (Sagi and Rokach, 2021; Romero Martínez et al., 2024). The complexity of the model, including many decision trees, makes it difficult to understand the specific contributions of individual features to the final output. In recent decades, explainable artificial intelligence (XAI) techniques, such as SHapley Additive exPlanations (SHAP), have been widely utilized to interpret the behaviour of machine learning models and identify factors affecting their predictions on unknown samples. Recent years have witnessed a growing academic interest in glass-box algorithms, offering enhanced interpretability of AI models. Unlike black-box algorithms, glass-box algorithms, such as Explainable Boosting Machine (EBM), do not require external tools to explain the contribution of features and provide direct interpretability. The EBM integrates the power of Generalized Additive Models and gradient boosting, focusing on model interpretability. Although EBM was developed as a competitor to robust algorithms such as Random Forest, LightGBM, or XGBoost, it has been employed in a limited number of studies, especially in remote sensing.

To address this gap in the literature, this study evaluated the potential application of EBM in determining the geographical distribution of poplar plantations using remotely sensed imagery having different resolutions (e.g., Sentinel-2 and PlanetScope images). In addition, the statistical and visual performance of the algorithm was compared by employing different spectral features with the XGBoost algorithm integrated with the SHAP.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-M-6-2025-227-2025

2. Study Area and Dataset

The Akyazi district of Sakarya Province, located in the northwestern part of Turkey, was chosen as the study area (Figure 1). According to 2007 records from the Sakarya Provincial Directorate of Agriculture, Akyazi stands out as Turkey's most significant poplar cultivation center, with around 2 km^2 of poplar plantation areas within its borders. In addition to poplar, maize and hazelnut species, which are predominantly cultivated, contribute to the variety of agricultural activities in the region.

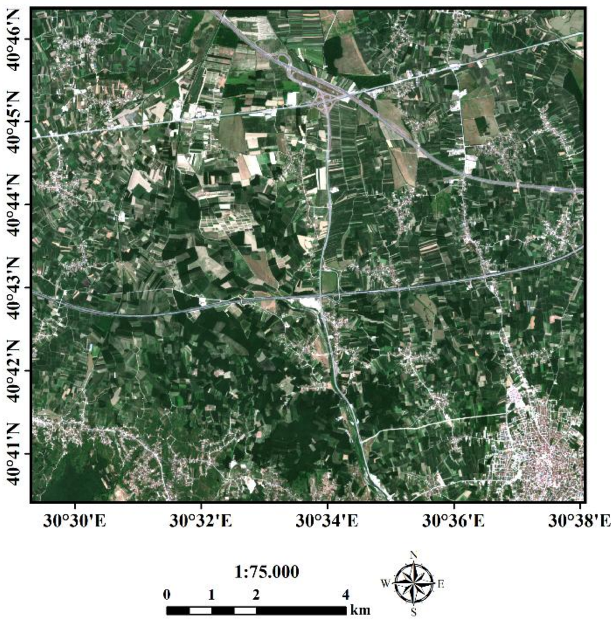

Figure 1. Location of the study area.

In this study, atmospherically corrected Sentinel-2 and PlanetScope images were used as the main data source. Sentinel-2 was developed by the European Space Agency, and it provides 13 spectral bands with spatial resolutions of 10m, 20m, and 60m in the visible, Red Edge, NIR and SWIR regions of spectrum. Meanwhile, PlanetScope images are commercially served by Planet company and were provided free of charge to utilize in this study as part of the Planet Education and Research Programme. It is an Earth Observation satellite providing 8 spectral bands (i.e., six in the visible region and two in the Red Edge and NIR regions) at a 3m spatial resolution with daily coverage. Note that Sentinel-2 and PlanetScope images acquired on the same date (2 July 2024) were utilized in the classification process to make an objective comparison.

3. Methodology

The methodology applied in this research involves the following process steps: i) image pre-processing, ii) image classification using EBM and XGBoost, and iii) quantitative and qualitative evaluation of results. In the first step, the three spectral bands of Sentinel-2 at 60m were excluded, and the remaining 10 bands were evaluated. The six bands at 20m spatial resolution were downscaled to 10m using the nearest neighbour resampling technique. Moreover, texture features including mean, variance, homogeneity, contrast, entropy, and correlation, were calculated for each spectral band of both satellite images and stacked with original spectral bands to enhance thematic mapping accuracy. Secondly, classification models of EBM and XGBoost were constructed and applied to the satellite imagery, producing two-class thematic maps representing poplar and non-poplar fields. In addition, the 15 most effective features for model predictions were identified, and models' performances in distinguishing poplar pixels were also analysed using these selected features. For this purpose, SHAP was applied to black-box XGBoost to interpret the model's internal behavior, and the most effective 15 features were selected for each imagery. In the last step, the confusion matrix-based overall accuracy (OA) and F-score (F-S) values were estimated for each considered dataset for accuracy assessment. In addition, all results were visually compared.

3.1 Explainable Boosting Machine (EBM)

The novel glass-box model, EBM, was engineered as a modern version of Generalized Additive Model (GAM), integrating both accuracy competitive with complex black-box models (e.g., Random Forest, CatBoost, and XGBoost) and interpretability and transparency (Nori et al., 2021). The training process including decision tree production, learning and gradient updates, operates on individual predictor feature at a time in round-robin cycle employing a minimal learning rate (Maxwell et al., 2021). A round-robin cycle is performed over features to avoid the effects of co-linearity, to learn the appropriate feature function designed for pairwise interactions and individual features, and to illustrate how each feature influences the model's estimation. Thereby, EBM's decision-making mechanism and final estimations can be interpreted globally and locally interpreted from both local and global perspectives (Liu and Sun, 2023).

3.2 Extreme Gradient Boosting (XGBoost)

One of the most robust black-box classification algorithms, XGBoost, was presented by Chen and Guestrin (2016) depending on gradient boosting theories. This algorithm iteratively train weak learners and integrate the results to contruct the efficient predictive model. The process begin equally weighting each sample in the training dataset and the first decision tree is trained using these samples. The subsequent iteration focuses on correcting the misidentified samples by increasing their weights (Georganos et al., 2018; Rumora et al., 2020).

3.3 Shapley Additive exPlanations (SHAP)

Despite demonstrating outstanding performance and accuracy across various remote sensing applications, the decision-making mechanisms of machine learning models lack interpretability and transparency, diminishing their trustworthiness. These models were hence referred to as black-box (Minh et al., 2022). To address challenges related to model opacity, Explainable Artificial Intelligence (XAI) have gained great renown. XAI focuses on generating visualizations and interactions for users to facilitate understanding of the impact of how the decision-making process impacts the outputs (Iban and Aksu, 2024). In this research, a prominent game-theoretic XAI tool, SHAP, was adopted to identify and interpret the most effective features of the XGBoost algorithm. To ensure interpretability, SHAP applies the additive feature imputation strategy representing the output model as a linear combination of input parameters (Abdollahi and Pradhan, 2021).

4. Results

To construct classification models using EBM and XGBoost, 5,000 samples representing poplar pixels and 10,000 samples corresponding to non-poplar pixels (i.e., maize, hazelnut, pasture, man-made structures, impervious surface, water, and other tree species) were collected as the training dataset for each satellite image, separately. The accuracy assessment of produced thematic maps was performed with 5,000 samples selected for each class (i.e., in total 10,000 samples).

4.1 Results of Sentinel-2 Image

The statistical results of EBM and XGBoost algorithms in poplar tree detection using Sentinel-2 image are given in Table 1. The results demonstrated that the EBM algorithm outperformed XGBoost in identifying poplar trees using Sentinel-2 imagery. In other words, XGBoost classified the Sentinel-2 image with an OA value of 92.15% using all features (i.e., spectral bands and texture features), while the F-S value representing poplar class accuracy was 92.65% in this map. On the contrary, EBM yielded about 1% greater overall and class-based accuracy values.

Features	Accuracy Metric	EBM	XGBoost
All	OA (%)	93.48	92.15
	F-S (%)	93.89	92.65
Selected features	OA (%)	94.27	92.84
	F-S (%)	94.56	92.91

Table 1. Accuracy assessment results for Sentinel-2 image.

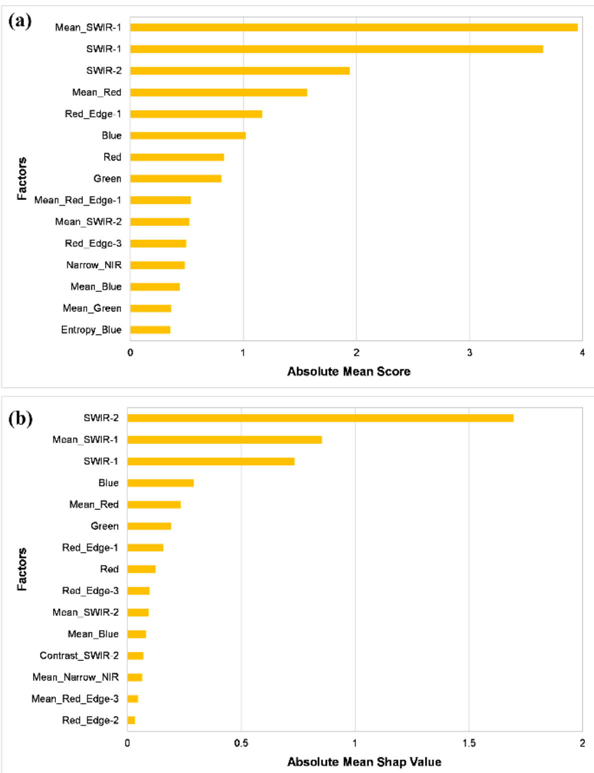

Figure 2. The 15 most influential factors determined by (a) EBM and (b) SHAP for Sentinel-2 imagery.

The classification performance of the models was also assessed based on the use of limited or selected spectral features. For this purpose, the absolute average scores of the 15 most effective factors in the model structure established by EBM and XGBoost were calculated for the poplar class, as shown in Figure 2. Figure 2a presents the absolute mean scores of the most influential factors identified by the EBM model from the spectral bands and derived texture features. On the other hand, Figure 2b shows the absolute mean Shap values for the most important features identified through SHAP analysis of the XGBoost model. Note that the higher the absolute mean value indicates the greater the influence on model prediction. Analysis of the graphs showed that the SWIR bands of Sentinel-2 (i.e., SWIR-1 and SWIR-2) and the mean texture feature of SWIR-1 were the most prominent features considered by both models for mapping poplar afforestation fields. These were followed by spectral values of visible bands (i.e., Blue, Green, and Red) and the Red Edge-1 band. In addition, within the calculated texture features, the mean was the most outstanding feature, while entropy and contrast indicated different influences on the predictions of EBM and XGBoost models. Moreover, it was observed that the NIR band of the Sentinel-2 image with a resolution of 10 m could not provide valuable spectral information for the discrimination of poplar afforestation fields. The thematic mapping process was also performed with selected features by both algorithms and the estimated accuracy values were given in Table 1. As can be seen, employing limited yet impactful features yielded superior accuracy values against the previous case of EBM and XGBoost. Additionally, the enhanced feature selection strategy employed by EBM contributed more significantly to improving map accuracy than that of XGBoost.

4.2 Results of PlanetScope Image

The classification performances of the EBM and XGBoost algorithms were also evaluated for the PlanetScope dataset. The estimated accuracy results are given in Table 2. As can be seen, a similar scenario was obtained with Sentinel-2 results in terms of algorithm performance comparisons, whereas significantly lower accuracy values (i.e., about 10%) were estimated. Analysis of the results revealed approximately 1% increase in OA and class-based accuracy (F-S) when using the EBM algorithm compared to XGBoost, utilizing the dataset containing the spectral bands and texture features.

Features	Accuracy Metric	EBM	XGBoost
All	OA (%)	83.92	82.71
	F-S (%)	83.14	81.75
Selected features	OA (%)	84.44	83.87
	F-S (%)	83.79	83.04

Table 2. Accuracy assessment results for PlanetScope image.

Additionally, feature selection strategy was implemented analyzing the visual charts of EBM and SHAP given in Figure 3. According to charts, while the spectral information derived applying the mean texture feature was recognised as among the most significant features for poplar class by both approaches, the texture feature of the Red Edge band was the most impactful band. On the other hand, the mean texture feature calculated from the yellow band was more prominent than spectral features derived from green channels (i.e., Green and Green-I) for XGBoost’s model structure. Employing presented 15 features, EBM and XGBoost models were rebuilt to produce poplar map and statistical results were given in Table 2. As can be seen, few impactful features for the classification process conducted by EBM and XGBoost presented better accuracy values. As with Sentinel-2, using selected features improved accuracy for both algorithms, with EBM producing the highest overall performance.

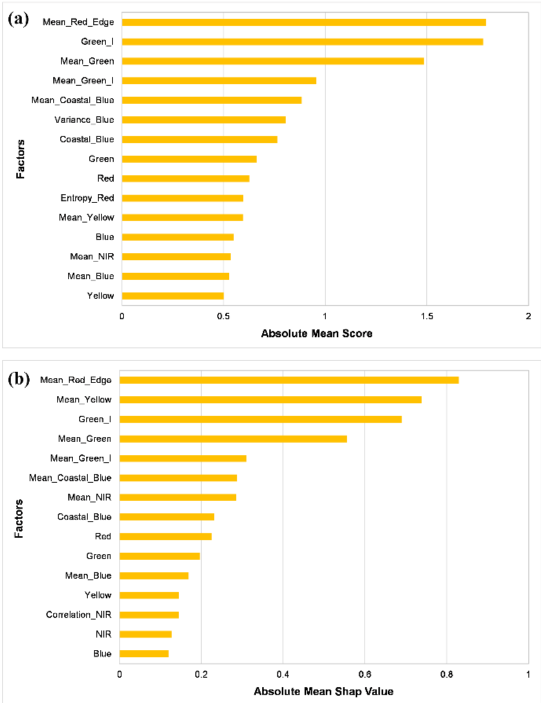

Figure 3. The 15 most influential factors determined by (a) EBM and (b) SHAP for PlanetScope imagery.

4.3 Visual Assessment

To evaluate the results visually, Figure 2 depicts the highest-accuracy thematic maps produced utilizing Sentinel-2 and PlanetScope. Additionally, the true- and false-colour Sentinel-2 images were presented as a base map to highlight the distribution of poplar plantations. Key differences between the visuals of Sentinel-2 and PlanetScope images were highlighted with red and white dashed circles. While poplar afforestation parcels are generally identified more accurately using Sentinel-2 satellite imagery, the misclassifications observed in PlanetScope data are considered to result from the spectral similarity between poplar and other vegetation types, such as forest and agricultural plants.

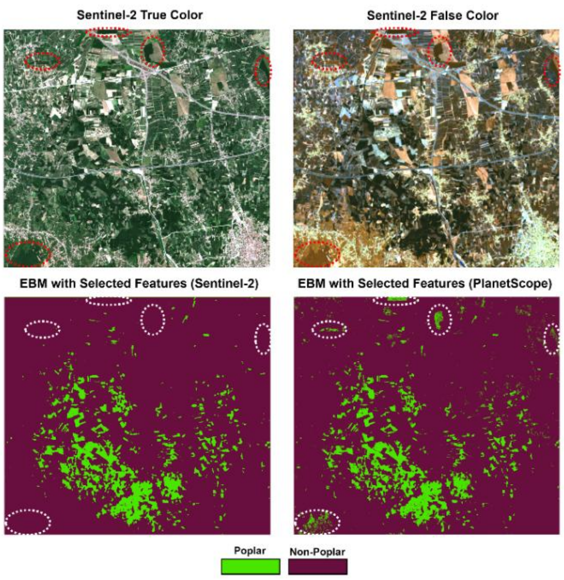

Figure 4. Visual results of Sentinel-2 and PlanetScope images.

5. Conclusions

In this research, the qualitative and quantitative performance of the glass-box EBM algorithm was evaluated in identifying poplar plantations over two satellite images with different resolutions. For this purpose, the classification process was conducted by considering the most effective features contributing to the classification within the dataset, determined by EBM, as well as the dataset combining spectral bands and texture features. The results were compared with those obtained using the black-box XGBoost algorithm combined with SHAP analysis. Several conclusions can be inferred from the comparative results. First, EBM presented superior mapping performance compared to XGBoost for all scenarios. It supports the hypothesis that EBM achieves accuracies comparable to cutting-edge algorithms. Second, poplar mapping resulted in significantly greater accuracies utilizing Sentinel-2 with higher spectral and lower spatial resolution versus PlanetScope. This indicates the relevance of spectral features including SWIR bands for accurately extraction of tree species such as poplar. Finally, the more satisfactory classification results were obtained by employing fewer yet impactful features. This highlights the importance of feature selection in enhancing model improvement, especially when using lower-resolution imagery.

References

Abdollahi, A., Pradhan, B., 2021. Urban Vegetation Mapping from Aerial Imagery Using Explainable AI (XAI). *Sensors*, 21, 4738. doi.org/10.3390/s21144738.

Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, NY, USA, 785–794. doi.org/10.1145/2939672.2939785.

Colkesen, I., Kavzoglu, T., Atesoglu, A., Tonbul, H., Ozturk, M.Y., 2023. Multi-seasonal evaluation of hybrid poplar (P. Deltoides) plantations using Worldview-3 imagery and State-Of-The-Art ensemble learning algorithms. *Adv. Sp. Res.*, 71, 3022–3044. doi.org/10.1016/j.asr.2022.10.044.

Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., Wolff, E., 2018. Very High Resolution Object-Based Land Use–Land Cover Urban Classification Using Extreme Gradient Boosting. *IEEE Geosci. Remote Sens. Lett.*, 15, 607–611. doi.org/10.1109/LGRS.2018.2803259.

Iban, M.C., Aksu, O., 2024. SHAP-Driven Explainable Artificial Intelligence Framework for Wildfire Susceptibility Mapping Using MODIS Active Fire Pixels: An In-Depth Interpretation of Contributing Factors in Izmir, Türkiye. *Remote Sens.*, 16, 2842. doi.org/10.3390/rs16152842.

Kavzoglu, T., Colkesen, I., Atesoglu, A., Tonbul, H., Yilmaz, E.O., Ozlusoylu, S., Ozturk, M.Y., 2024. Construction and implementation of a poplar spectral library based on phenological stages for land cover classification using high-resolution satellite images. *Int. J. Remote Sens.*, 45, 2049–2072. doi.org/10.1080/01431161.2024.2326800.

Khayatinejad Aqgonbad, S., Hossein Zadeh, O., Hajjarian, M., Abdi, M.R., 2022. Impact of Key Stakeholders’ Participation in Strategic Decision-Making Regarding Poplar Tree Farming Development. *Small-scale For.*, 21, 229–247. doi.org/10.1007/s11842-021-09496-3.

Kocer, S., Kara, M.S., 2021. Economics of poplar (clones of P.x euramericana ve P. deltoides) plantations established in closer spacings. *Turkish J. For.*, 22, 257–270. doi.org/10.18182/tjf.949462.

Liu, G., Sun, B., 2023. Concrete compressive strength prediction using an explainable boosting machine model. *Case Stud. Constr. Mater.*, 18, e01845. doi.org/10.1016/j.cscm.2023.e01845.

Maxwell, A.E., Sharma, M., Donaldson, K.A., 2021. Explainable Boosting Machines for Slope Failure Spatial Predictive Modeling. *Remote Sens.*, 13, 4991. doi.org/10.3390/rs13244991.

Minh, D., Wang, H.X., Li, Y.F., Nguyen, T.N., 2022. Explainable artificial intelligence: a comprehensive review. *Artif. Intell. Rev.*, 55, 3503–3568. doi.org/10.1007/s10462-021-10088-y.

Nori, H., Caruana, R., Bu, Z., Shen, J.H., Kulkarni, J., 2021. Accuracy, Interpretability, and Differential Privacy via Explainable Boosting. Meila, M., Zhang, T. (Eds.), *Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research*. PMLR, 8227–8237..

Romero Martínez, M., Pozuelo Campillo, J., Carmona Ibáñez, P., 2024. Ethical transparency in business failure prediction: uncovering the black box of xgboost algorithm. *Spanish J. Financ. Account. / Rev. Española Financ. y Contab.*, 1–31. doi.org/10.1080/02102412.2024.2423498.

Rumora, L., Miler, M., Medak, D., 2020. Impact of Various Atmospheric Corrections on Sentinel-2 Land Cover Classification Accuracy Using Machine Learning Classifiers. *ISPRS Int. J. Geo-Information*, 9, 277. doi.org/10.3390/ijgi9040277.

Sagi, O., Rokach, L., 2021. Approximating XGBoost with an interpretable decision tree. *Inf. Sci. (Ny).*, 572, 522–542. doi.org/10.1016/j.ins.2021.05.055.

Velioglu, E., Akgul, S., 2016. Poplars and willows in Turkey: Country progress report of the national poplar commission time period: 2012–2015, poplar and fast growing forest trees.