ABSTRACT

Carbon stocks are critical to climate change mitigation by capturing atmospheric carbon and storing it in biomass. However, carbon stock estimation faces challenges due to data complexity and the need for efficient analytical methods. This study introduces a carbon stock estimation method that integrates the XGBoost algorithm with VGG16 feature extraction and feature selection techniques to analyze GEE and Drone image datasets. The model is evaluated through four scenarios: without feature selection, using Information Gain, using Feature Importance, and using Recursive Feature Elimination. These scenarios aim to compare feature selection methods to identify the best one for processing complex environmental data. The experimental results show that RFE significantly outperforms other methods, achieving an average RMSE of 6651.62, MAE of 2297.57, and R^2 of 0.7673. These findings underscore the importance of feature selection in optimizing model performance, particularly for high-dimensional environmental datasets. RFE shows superior accuracy and efficiency by retaining the most relevant features but requires more computational resources. For applications that prioritize time and resource efficiency, Information Gain or Feature Importance can serve as a practical alternative with slightly reduced accuracy. This research highlights the value of integrating feature selection techniques into machine learning models for environmental data analysis. Future research could explore alternative feature extraction methods, combine RFE with other approaches, or apply advanced techniques such as Boruta or genetic algorithms. These efforts will further refine carbon stock estimation models, paving the way for broader applications in environmental data analysis.