Researches on Advances in Forest Ecosystem Monitoring

Editors

Gonca Ece ÖZCAN

Korhan ENEZ

BİDGE

BIDGE Publications

Researches on Advances in Forest Ecosystem Monitoring

Editors: Gonca Ece ÖZCAN & Korhan ENEZ

ISBN: 978-625-372-161-9

Page Layout: Gözde YÜCEL

1st Edition:

Publication Date: 30.11.2024

BIDGE Publications,

Certificate No: 71374

www.bidgeyayinlari.com.tr - bidgeyayinlari@gmail.com

Krc Bilişim Ticaret ve Organizasyon Ltd. Şti.

Güzeltepe Mahallesi Abidin Daver Sokak Sefer Apartmanı No: 7/9 Çankaya / Ankara

PREFACE

Forest ecosystems are one of the most important components of the world's ecosystems. These ecosystems, which play a critical role in both the conservation of biodiversity and the fight against climate change, are increasingly attracting attention with their sustainable management and monitoring requirements. In recent years, the methods used to understand the health and functionality of forest ecosystems have undergone a significant transformation with technological advancements and innovations in data analysis. Remote sensing, artificial intelligence, deep learning, time series, and advanced modeling techniques have made this transformation possible.

"***Researches on Advances in Forest Ecosystem Monitoring***" brings together the latest research and applications in the monitoring and management of forest ecosystems. The book addresses various scientific studies on critical topics such as biomass and carbon estimation in forest ecosystems, detection of forest fires, time series analysis in the wood products industry, and controlling forest pests. Additionally, the innovative methods and technologies used under each topic ensure the more effective monitoring and management of forest ecosystems.

The first section addresses carbon estimation using the MARS modeling technique, while the second section details biomass predictions in pure Scots pine stands using dummy variable regression analysis. The third section addresses the detection of forest fires using the YOLO algorithm from digital forest images. The fourth and fifth chapters include the evaluation of forest pests using remote sensing techniques. Finally, the sixth chapter explains the use of time series analysis in the wood products industry.

We hope that this book will be an important resource for scientists researching forest ecosystems, forest engineers, and nature and environmental experts. Additionally, it offers a guiding study for decision-makers and technology developers in forest management. The book addresses the challenges encountered in forest monitoring and management while also showcasing the potential of innovative technologies to overcome them.

Every step taken to understand better, protect, and manage forest ecosystems is of immense importance for future generations. We hope that this book will serve as a scientific resource contributing to the sustainability of forest ecosystems and helping make more effective decisions in forest management worldwide.

Editors

Gonca Ece ÖZCAN

Korhan ENEZ

İçindekiler

CHAPTER I

Predicting Aboveground Biomass using Multivariate Adaptive Regression Splines and Stepwise Regression

Alkan GÜNLÜ[1]
Semih KUTER[2]

Introduction

Forest ecosystems play a critical role in regulating the global climate by managing carbon emissions, absorbing carbon, and controlling energy and water cycles (Herold et al., 2019; Puliti et al., 2021). Forest biomass is a key component for understanding and predicting the global carbon cycle, essential for effective forest management (Moradi et al., 2022; Santoro et al., 2021). Accurately estimating forest biomass is essential for effective sustainable forest management (Khan et al., 2020; Maack et al., 2015).

[1] Prof. Dr. Alkan GÜNLÜ, Çankırı Karatekin Üniversitesi, Orman Fakültesi, Orman Mühendisliği Bölümü, Çankırı/Türkiye, Orcid: 0000-0003-4759-3125, alkangunlu@karatekin.edu.tr,

[2] Doç. Dr. Semih KUTER, Çankırı Karatekin Üniversitesi, Orman Fakültesi, Orman Mühendisliği Bölümü, Çankırı/Türkiye, Orcid: 0000-0002-4760-3816, semihkuter@karatekin.edu.tr

Total forest biomass consists of two parts: below-ground and above-ground biomass (Tetemke et al., 2024; Vicharnakorn et al., 2014). In the literature, studies mostly focus on estimating above-ground biomass (Askar et al., 2018; Chunhua Li et al., 2021; Wang et al., 2024). Above-ground biomass (AGB) is accurately estimated using the allometric relationship between diameter at breast height and tree height, measured through field surveys. However, applying this method to large forested areas is challenging, costly, and time-consuming (Lu, 2006).

In contrast, with the development of remote sensing technologies over the past twenty years, studies have combined remote sensing data with field measurements to estimate AGB at a regional level (Behera et al., 2024; Goetz et al., 2009; Günlü et al., 2014; Vafaei et al., 2018). The literature in this area shows that regression analysis is commonly used to predict relationships between variables obtained from remote sensing data and AGB measured in the field (Günlü et al., 2014; Chao Li et al., 2019; Lu, 2006).

A regression model can be developed using variables from field data and remote sensing data. However, traditional regression modeling techniques cannot effectively capture the complex, non-linear relationships between AGB and satellite data. Therefore, to improve the model's ability to predict AGB non-linearly, machine learning methods such as decision trees, K-nearest neighbors (KNN), artificial neural networks (ANN), support vector machines (SVM), and multivariate adaptive regression splines (MARS) are used to model the relationships between ground-measured biomass and variables obtained from remote sensing data (Migolet et al., 2022; Nelson et al., 2009; Vaglio Laurin et al., 2016).

In this study, the relationships between AGB values calculated from field measurements in pure Scots pine stands and the Sentinel-2 band reflectance values were examined using multiple regression and MARS modeling techniques.

Material and Methods

Study Area

The study area (see Figure 1) is located in the Erenlerkös Forest Planning Unit within the boundaries of the Kargı State Forest Enterprise, the Amasya Regional Directorate of Forestry (40° 58’ 06" – 41° 07’ 01" N, 34° 19’ 22" – 34° 41’ 09" E) in Türkiye. Elevations in the study area range from 300 m to 2096 m. The average annual precipitation is 477.5 mm, and the mean temperature is 11.2°C (Anonymous, 20218). The study area covers a total of 21,360.8 ha, of which 16,100.5 ha is forested. Dominant tree species include *Pinus sylvestris* L., *Pinus brutia* Ten., *Pinus nigra* Arnold., *Carpinus betulus* L., *Juniperus* spp., and *Quercus frainetto*. The specific focus of this study is a 2,147.6 ha pure Scots pine forest within this unit.

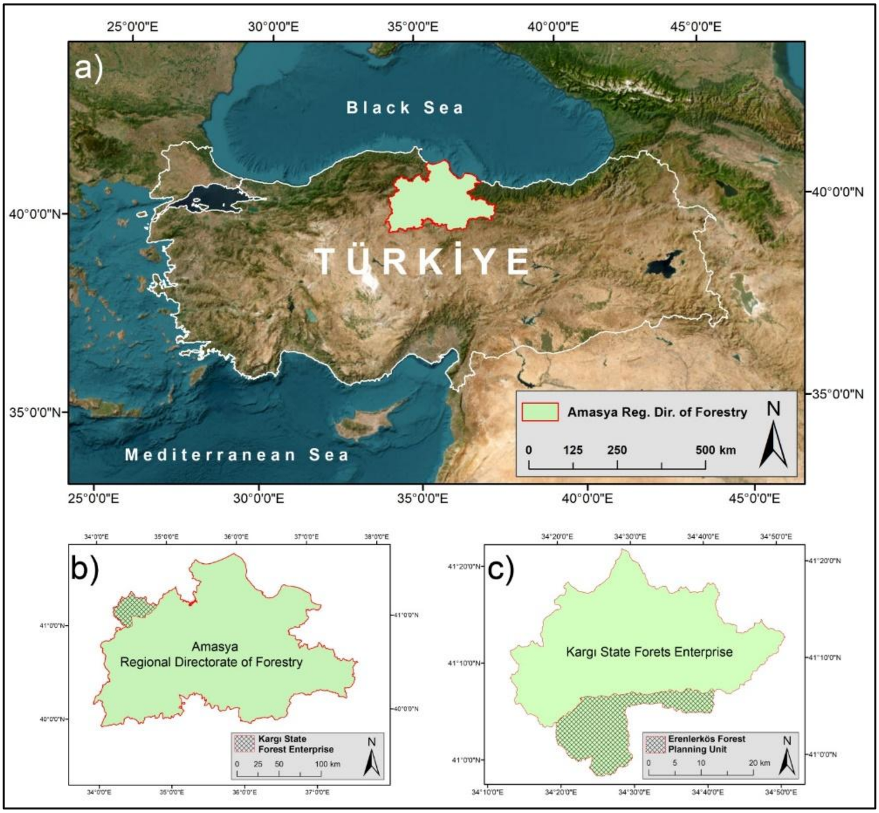

Figure 1: Location of study area

Data

Ground Measurements

The study used data from a total of 195 sample plots (see Figure 2). For each plot, the aboveground biomass (AGB) of individual trees, measured in kilograms, was calculated based on the diameter at breast height ($d_{1.3}$) data from the inventory. The single-entry AGB equation developed by Yavuz et al. (2010) (Equation 1) was applied to estimate the AGB of each tree. The total AGB for each plot was then obtained by summing the individual tree AGB values. Finally, AGB per hectare was calculated for each plot, based on its area (400, 600, or 800 m²).

$$AGB = 6.89952 - 2.423793 \times d - 0.373438 \times d^2 \quad (1)$$

where AGB is aboveground biomass and d indicates diameter at breast height (cm).

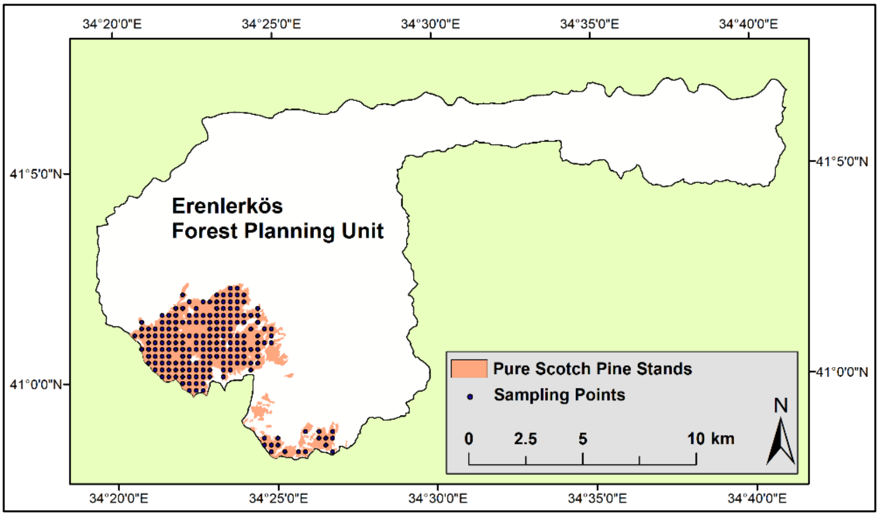

Figure 2: Locations of the sample plots in Erenlerkös planning unit

Remote sensing data

A Sentinel-2 image acquired on 12th of September, 2018 was downloaded from the Sentinel Hub website (https://apps.sentinel-hub.com/). The Sentinel-2 instrument has 13 spectral bands, with 4 bands at 10 m resolution, 6 bands at 20 m resolution, and 3 bands at 60 m resolution. In this study, 10 of these bands, specifically those with 10 and 20 m spatial resolution, were selected for analysis. Details about Sentinel-2's spectral channels are presented in Table 1. The image was calibrated in QGIS Desktop 3.8.1 to produce reflectance images for the 10 selected bands. Next, using ArcGIS Desktop 10.6.1, a buffer zone was applied around each band based on sample plot sizes, with radii of 15.96 m for low-coverage plots, 13.82 m for medium-coverage plots, and 11.28 m for full-coverage plots. The "*zonal statistics*" tool in ArcGIS Desktop 10.6.1 was then used to compute the minimum, maximum, mean, sum, and range of reflectance values for each sample plot.

Table 1: Sentinel-2 band designations

Bands	**Wavelength Interval (nm)**	**Spatial Resolotion (m)**
Band 2	458 – 523	10
Band 3	543 – 578	10
Band 4	650 – 680	10
Band 5	698 – 713	20
Band 6	733 – 748	20
Band 7	773 – 793	20
Band 8	785 – 900	10
Band 8A	855 – 875	20
Band 11	1565 – 1655	20
Band 12	2100 – 2280	20

Statistical Analysis

Multiple regression and MARS modeling techniques were used as statistical analyses in the study.

Multiple Regression Analysis

The multiple regression model structure used in the study is given in the equation below.

$$AGB = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n + \varepsilon \qquad (2)$$

In the equation; the AGB: aboveground biomass, β_0, β_1, β_2,…, β_n are the equation parameters, X_1, X_2, X_3, …, X_n is the minimum, maximum, average, total, and range reflectance values obtained for each sample plot for the Sentinel-2 image, and ε is the error term.

Multivariate Adaptive Regression Splines

Multivariate adaptive regression splines (MARS) is a regression method used to model complex relationships between input variables and an output variable (Friedman, 1991). It is especially useful when the relationship between variables is non-linear or when variables interact with each other, meaning it is effective in cases where different variables influence one another.

MARS is based on the idea of piecewise linear regression. Instead of fitting a single equation to the data, MARS divides the data into different regions and fits simple linear regressions (called “basis functions”) to each region. These regions are defined by “knots” where the data is split. The best knot locations are chosen automatically during model building. Initially, MARS creates many candidate basis functions. These are simple linear functions applied to subsets of the data, generally starting as piecewise linear segments. The model first fits all possible basis functions to the data and selects the ones that best improve the fit (Hastie et al., 2009).

- Forward Step: The algorithm adds one basis function at a time to improve the model, continuing until the model fits the data well (though it may risk overfitting).

- Backward Step: To prevent overfitting, MARS uses a backward pruning step that removes basis functions that don't significantly improve accuracy, aiming to keep the model simple while capturing key patterns in the data.

- Basis Functions: These are simple functions that "bend" at the knots to capture changes in the data. They can take forms like $(x\text{-}c)_+$ or $(c\text{-}x)_+$, where x is the input variable and c is the knot location.

- Interaction Terms: MARS can also model interactions between variables, meaning it captures cases where the effect of one variable depends on the value of another.

- Selecting Knots and Variables: During the forward pass, MARS decides where to place knots and which variables to include based on how well they fit the model, while also considering a cost function that penalizes added complexity.

The MARS model uses several key parameters to perform well on data. These parameters control model complexity, interaction degree, and overall performance:

- max_degree: This parameter sets the maximum degree of interaction between variables in the model. If max_degree = 1, the model uses only individual variables (no interaction terms). Higher values allow more complex interactions, which makes the model more expressive but potentially harder to interpret. Very high values may lead to overfitting.

- penalty: MARS uses a penalty term to control model complexity, applying a penalty for each additional basis function. A higher penalty leads to a simpler model using fewer basis functions, which helps prevent overfitting but can also make the model fit the data less well.

- thresh (Threshold Value): This parameter sets the error improvement threshold for adding a new basis function during the forward pass. If a new basis function improves the error below this threshold, it's accepted. A low threshold allows many functions to be added, making the model complex and potentially overfit. A high threshold adds only functions that provide major improvements, leading to a simpler model.

- Max Number of Terms: This limits the number of basis functions in the model, checked during the forward pass to keep the model from becoming overly complex. A lower max term count keeps the model simple, while a higher count allows a more complex, better-fitting model but with a risk of overfitting.

Balancing complexity and overfitting is essential when adjusting a MARS model. These parameters help control the number of knots, basis functions, and interaction levels, ensuring the model remains flexible yet general. Proper parameter tuning aids in creating a model that explains the data well without overfitting.

Advantages of MARS:

- Captures non-linear relationships: By using piecewise linear functions, MARS can model complex, non-linear patterns that traditional linear models may miss.
- Automatic feature selection: MARS automatically selects important input variables, reducing the need for manual feature engineering.
- Handles interactions: MARS can capture variable interactions without needing them to be specified in advance.

Limitations:

- Prone to overfitting: Without proper adjustments, MARS can fit too many basis functions, causing overfitting to the training data. The backward step reduces this, but careful tuning is essential.
- Less interpretable than simple models: While more flexible than linear regression, MARS models are harder to interpret due to their piecewise and interaction-based structure.

MARS is commonly used in regression problems in fields like environmental modeling, finance, and engineering, especially when relationships between variables are complex and non-linear. It's a good choice when interactions among variables play a key role in explaining the outcome since it can automatically identify and model these interactions.

In summary, MARS is a powerful prediction tool, especially when relationships between variables are non-linear or interactive. It adapts to the data, making it a flexible method that can improve accuracy in many real-world applications. However, careful adjustments are needed to avoid overfitting and keep the model understandable.

Performance metrics

Various criteria were applied to assess the prediction accuracy of the modeling techniques. These criteria include the correlation coefficient (R), coefficient of determination (R^2), root-mean-square error (RMSE), and mean absolute error (MAE).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \qquad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \qquad (4)$$

$$R = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{y_i - \overline{y}}{s(y)}\right)\left(\frac{\hat{y}_i - \overline{\hat{y}}}{s(\hat{y})}\right) \qquad (5)$$

$$R^2 = 1 - \frac{(N-1)\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{(N-p)\sum_{i=1}^{N}(y_i - \bar{y})^2} \qquad (6)$$

where N is the number of observations, p is the number of parameters, y_i are measured values of the AGB, $\hat{y}_i$ are predicted AGB, $\bar{y}$ is a mean of measured AGB values, $s(y)$ and $s(\hat{y})$ are the standard deviations of the observed and the predicted AGB values, respectively.

Results and Discussion

In the study, the relationships between variables derived from Sentinel-2 satellite imagery and AGB data obtained from ground measurements were modeled using MARS and MLR techniques. The model and test performances are presented in Table 2. Additionally, the observation and prediction graphs for the model (i.e., training) and test datasets, according to both modeling techniques, are shown in Figures 3-6.

Table 2: The performances of MARS and MLR on training and test datasets

Method	Training			Test		
	RMSE	MAE	*R*	RMSE	MAE	*R*
MLR	62.1685	46.9446	0.4603	69.0457	55.0238	0.3170
MARS	47.1735	37.3444	0.7419	49.2146	41.1786	0.7209

As indicated by the statistical performance metrics in Table 2, for the MLR model, the training RMSE was 62.17, and the MAE was 46.94, with a correlation coefficient (*R*) of 0.4603. On the test dataset, the model's performance decreased, with an RMSE of 69.05, an MAE of 55.02, and a lower *R* value of 0.3170. These results indicate that the MLR model had relatively poor predictive power, especially when applied to unseen data, showing limited generalization ability.

In contrast, the MARS model demonstrated significantly better performance. On the training dataset, the RMSE was 47.17, and the MAE was 37.34, with a much higher *R* value of 0.7419. Similarly, the MARS model maintained strong predictive accuracy on the test dataset, with an RMSE of 49.21, an MAE of 41.18, and an *R* value of 0.7209. These results suggest that the MARS model not only fit the training data well but also generalized effectively to the test data, outperforming the MLR model in all metrics.

The MLR and MARS models obtained have different sets of predictor variables. MLR uses only two predictor variables, whereas the final MARS model employs 4 predictors, as given in Table 3. Only one predictor variable is common in both models, i.e., B7_Sum.

Table 3. The predictor variables involved in the final MLR and MARS models

Model	Selected Predictor Variable Set
MLR	B2_Sum, B7_Sum
MARS	B5_Mean, B7_Sum, B7_Max, B12_Range

Overall, the MARS model proved to be a more robust and accurate method for predicting AGB, providing better results in both training and test datasets compared to the MLR model. This highlights the suitability of the MARS approach for capturing complex relationships in the data.

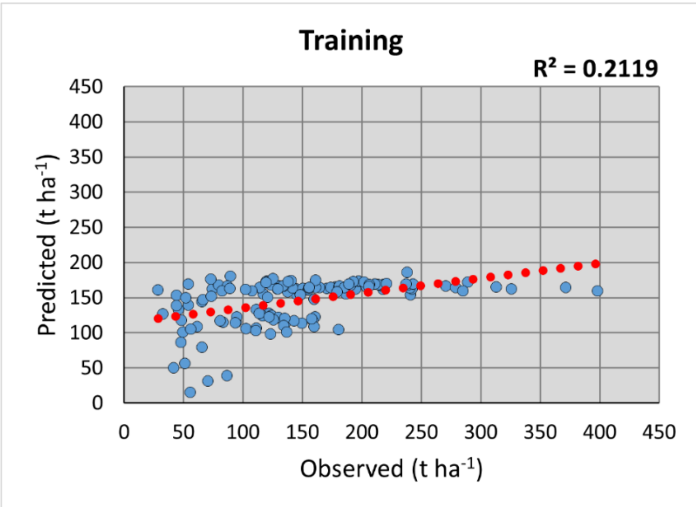

Figure 3: The graph of observed vs predicted values for MLR on training

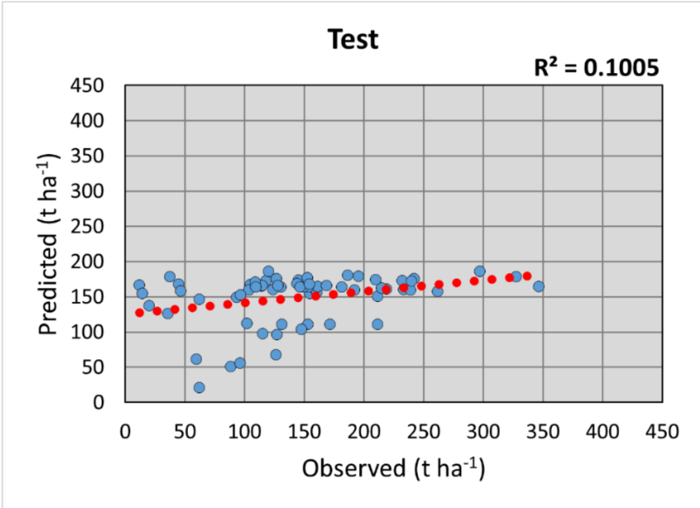

Figure 4: The graph of observed vs predicted values for MLR on testing

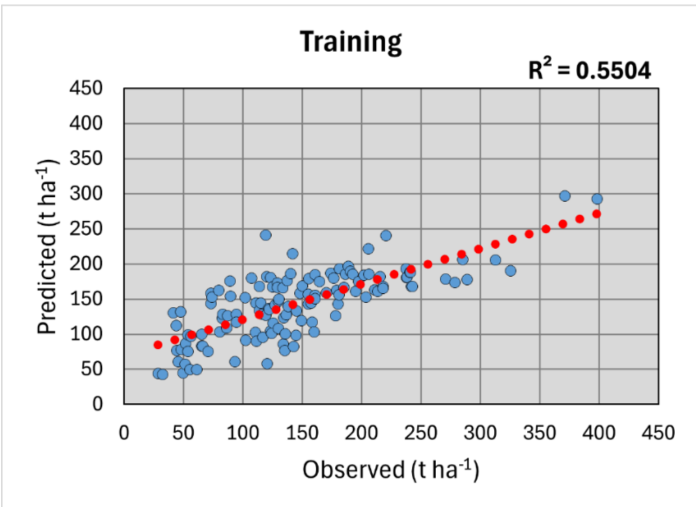

Figure 5: The graph of observed vs predicted values for MARS on training

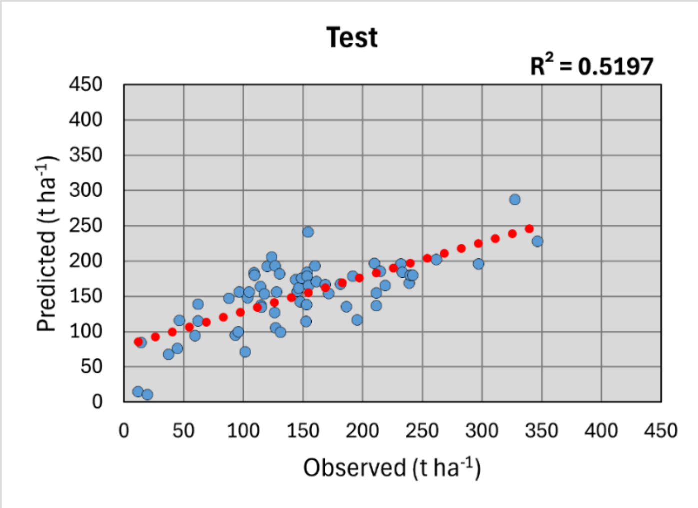

Figure 6: The graph of observed vs predicted values for MARS on testing

When examining several related studies in the literature, Günlü et al. (2014) modeled the relationships between band reflectance values derived from Landsat TM satellite imagery and AGB using the MLR method in pure black pine stands. The coefficient of determination (R^2 = 0.465) obtained from their model was higher compared to our study (R^2 = 0.2119). Similarly, in a study by Turgut & Günlü (2022), the relationships between band brightness values from Landsat 8 OLI satellite imagery and AGB were modeled using the MLR method in pure black pine stands, and their coefficient of determination (R^2 = 0.445) was also higher than ours.

In a study by Demirel et al. (2023), the relationships between band reflectance values from Sentinel-2 satellite imagery and aboveground carbon in pure cedar stands were modeled using the

MLR method, with a coefficient of determination of R^2 = 0.456. In another study by Bulut (2023), the relationships between AGB and band reflectance values from Landsat 8 OLI in pure Calabrian pine stands were modeled using MLR and support vector machine (SVM) techniques. The results indicated that R^2 was 0.50 for MLR and 0.67 for SVM.

In a study by Ayushi et al. (2024), the relationships between variables related to bands and vegetation indices from Sentinel-2 satellite imagery and AGB were predicted using different machine learning techniques, including SVM, MARS, k-NN, random forest, ANN, gradient boosting, and penalized regression. The model results showed that the coefficients of determination ranged from 0.40 to 0.65, with the MARS technique yielding a coefficient of determination of 0.59, similar to the results obtained in our study.

Lastly, in a study by Safari et al. (2017), conducted in two different oak forests (undegraded and highly degraded), the relationships between variables derived from Landsat 8 OLI bands and AGB were analyzed using SVM, boosted regression trees (BRT), RF, and MARS techniques. For the MARS method, the coefficients of determination were 0.34 in highly degraded forests and 0.56 in undegraded forests.

Conclusion

This study evaluated the performance of MLR and MARS modeling techniques in predicting AGB using variables derived from 10- and 20-meter spatial resolution bands of Sentinel-2 satellite imagery. The results clearly indicate that the MARS model outperformed the MLR model in terms of accuracy and generalization. While the MLR model exhibited moderate predictive capability with relatively higher error rates, the MARS model demonstrated superior performance, with lower RMSE and MAE values and a significantly stronger correlation between observed and predicted AGB values.

The findings suggest that the MARS modeling technique is better suited for capturing the complex, non-linear relationships between AGB and variables obtained from satellite images, making it a more effective tool for biomass estimation. Given the limited use of MARS in both national and international studies, particularly in forest ecosystems, this study underscores the need to expand its application across different forest types. This would enhance the accuracy of AGB predictions and contribute to improved forest management and carbon stock assessments.

Future studies should explore the potential of combining MARS with other advanced machine learning techniques and increasing the quantity and diversity of data to further improve model performance. Additionally, applying MARS in different forest ecosystems within the country would provide a more comprehensive understanding of its capabilities and limitations in biomass estimation. One potential improvement could be increasing the amount of data used in the training of the model, which may further enhance prediction accuracy.

Additionally, in our country, there has been no study focusing on the prediction of AGB using the MARS modeling technique. Even in international literature, studies using the MARS technique to predict AGB are relatively scarce. Therefore, there is a need to increase the number of such studies, particularly focusing on the diverse forest ecosystems in our country. This would not only contribute to filling a gap in the national research landscape but also provide valuable insights into the application of MARS for AGB estimation across various ecological contexts.

Acknowledgement

We would like to express our sincere gratitude to the General Directorate of Forestry, Department of Forest Management and Planning, for providing the inventory data used in this study.

References

Anonymous, 2018. Amasya Orman Bölge Müdürlüğü, Kargı Orman İşletme Müdürlüğü, Erenlerkös Orman İşletme Şefliği, Fonksiyonel Orman Amenajman Planı, Orman Genel Müdürlüğü, Orman İdaresi ve Planlama Dairesi, Ankara.

Askar, Nuthammachot, N., Phairuang, W., Wicaksono, P. & Sayektiningsih, T. 2018. Estimating Aboveground Biomass on Private Forest Using Sentinel-2 Imagery. Journal of Sensors, 2018(1).

Ayushi, K., Babu, K. N., Ayyappan, N., Nair, J. R., Kakkara, A. & Reddy, C. S. 2024. A comparative analysis of machine learning techniques for aboveground biomass estimation: A case study of the Western Ghats, India. Ecological Informatics, 80, 102479.

Behera, D., Kumar, V. A., Rao, J. P., Padal, S. B., Ayyappan, N. & Reddy, C. S. 2024. Estimating Aboveground Biomass of a Regional Forest Landscape by Integrating Textural and Spectral Variables of Sentinel-2 Along with Ancillary Data. Journal of the Indian Society of Remote Sensing, 52(4), 917-929.

Bulut, S. 2023. Machine learning prediction of above-ground biomass in pure Calabrian pine (*Pinus brutia* Ten.) stands of the Mediterranean region, Türkiye. Ecological Informatics, 74, 101951.

Demirel, D., Günlü, A. & Sakıcı, O. E. 2023. Estimating Above-Ground Carbon of Taurus Cedar Stands Using Sentinel-2 Satellite Image: A Case Study of Elmalı Forest Enterprise. In 3rd International Congress on Engineering and Life Science. Trabzon.

Friedman, J. H. 1991. Multivariate adaptive regression splines. The Annals of Statistics, 19(1), 1-67.

Goetz, S. J., Baccini, A., Laporte, N. T., Johns, T., Walker, W., Kellndorfer, J., Houghton, R.A.& Sun, M. 2009. Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. Carbon Balance and Management, 4(1), 2. https://doi.org/10.1186/1750-0680-4-2

Günlü, A., Ercanli, I., Başkent, E. Z. & Çakır, G. 2014. Estimating aboveground biomass using Landsat TM imagery: A case study of Anatolian Crimean pine forests in Turkey. Annals of Forest Research, 57(2), 289-298.

Hastie, T., Tibshirani, R. & Friedman, J. 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). NY, USA: Springer.

Herold, M., Carter, S., Avitabile, V., Espejo, A. ., Jonckheere, I., Lucas, R., McRoberts, R. E., Næsset, E., Nightingale, J., Petersen, R., Reiche, J., Romijn, E., Rosenqvist, A., Rozendaal, D. M. A., Seifert, F. M., Sanz, M. J. & De Sy, V. 2019. The Role and Need for Space-Based Forest Biomass-Related Measurements in Environmental Management and Policy. Surveys in Geophysics, 40(4), 757-778.

Khan, M. R., Khan, I. A., Baig, M. H. A., Liu, Z.J. & Ashraf, M. I. 2020. Exploring the potential of Sentinel-2A satellite data for aboveground biomass estimation in fragmented Himalayan subtropical pine forest. Journal of Mountain Science, 17(12), 2880-2896.

Li, C., Li, Y. & Li, M. 2019. Improving Forest Aboveground Biomass (AGB) Estimation by Incorporating Crown Density and Using Landsat 8 OLI Images of a Subtropical Forest in Western Hunan in Central China. Forests, 10(2), 104.

Li, C., Zhou, L. & Xu, W. 2021. Estimating Aboveground Biomass Using Sentinel-2 MSI Data and Ensemble Algorithms for Grassland in the Shengjin Lake Wetland, China. Remote Sensing, 13(8), 1595.

Lu, D. 2006. The potential and challenge of remote sensing-based biomass estimation. International Journal of Remote Sensing, 27(7), 1297-1328.

Maack, J., Kattenborn, T., Fassnacht, F. E., Enßle, F., Hernández, J., Corvalán, P. & Koch, B. 2015. Modeling forest biomass using Very-High-Resolution data—Combining textural, spectral and photogrammetric predictors derived from spaceborne stereo images. European Journal of Remote Sensing, 48(1), 245-261.

Migolet, P., Goïta, K., Pambo, A.F.K. & Mambimba, A. N. 2022. Estimation of the total dry aboveground biomass in the tropical forests of Congo Basin using optical, LiDAR, and radar data. GIScience & Remote Sensing, 59(1), 431-460.

Moradi, F., Sadeghi, S. M. M., Heidarlou, H. B., Deljouei, A., Boshkar, E. & Borz, S. A. 2022. Above-ground biomass estimation in a Mediterranean sparse coppice oak forest using Sentinel-2 data. Annals of Forest Research, 65(1), 165-182.

Nelson, R., Ranson, K. J., Sun, G., Kimes, D. S., Kharuk, V. & Montesano, P. 2009. Estimating Siberian timber volume using MODIS and ICESat/GLAS. Remote Sensing of Environment, 113(3), 691-701.

Puliti, S., Breidenbach, J., Schumacher, J., Hauglin, M., Klingenberg, T.F. &Astrup, R. 2021. Above-ground biomass change estimation using national forest inventory data with Sentinel-2 and Landsat. Remote Sensing of Environment, 265, 112644.

Safari, A., Sohrabi, H., Powell, S. & Shataee, S. 2017. A comparative assessment of multi-temporal Landsat 8 and machine learning algorithms for estimating aboveground carbon stock in coppice oak forests. International Journal of Remote Sensing, 38(22), 6407-6432.

Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D. M. A., Avitabile, V., Araza, A., de Bruin, S., Herold, M., Quegan, S., Rodríguez-Veiga, P., Balzter, H., Carreiras, J., Schepaschenko, D., Korets, M., Shimada, M., Itoh, T., Moreno Martínez, Á., Cavlovic, J., Cazzolla Gatti, R., da Conceição Bispo, P., Dewnath, N., Labrière, N., Liang, J., Lindsell, J., Mitchard, E. T. A., Morel, A., Pacheco Pascagaza, A. M., Ryan, C. M., Slik, F., Vaglio Laurin, G., Verbeeck, H., Wijaya, A. & Willcock, S. 2021. The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations. Earth Syst. Sci. Data, 13(8), 3927-3950.

Tetemke, B. A., Birhane, E., Rannestad, M. M. & Eid, T. 2024. Aboveground biomass estimation and mapping using Sentinel-2 data in a dry afromontane forest. International Journal of Remote Sensing, 1-19.

Turgut, R.& Günlü, A. 2022. Estimating aboveground biomass using Landsat 8 OLI satellite image in pure Crimean pine (*Pinus nigra* JF Arnold subsp. *pallasiana* (Lamb.) Holmboe) stands: a case from Turkey. Geocarto International, 37(3), 720-734.

Vafaei, S., Soosani, J., Adeli, K., Fadaei, H., Naghavi, H., Pham, T. D. & Tien Bui, D. 2018. Improving Accuracy Estimation of Forest Aboveground Biomass Based on Incorporation of ALOS-

2 PALSAR-2 and Sentinel-2A Imagery and Machine Learning: A Case Study of the Hyrcanian Forest Area (Iran). Remote Sensing, 10(2), 172.

Vaglio Laurin, G., Puletti, N., Chen, Q., Corona, P., Papale, D. &Valentini, R. 2016. Aboveground biomass and tree species richness estimation with airborne lidar in tropical Ghana forests. International Journal of Applied Earth Observation and Geoinformation, 52, 371-379.

Vicharnakorn, P., Shrestha, R. P., Nagai, M., Salam, A. P. & Kiratiprayoon, S. 2014. Carbon Stock Assessment Using Remote Sensing and Forest Inventory Data in Savannakhet, Lao PDR. Remote Sensing, 6(6), 5452-5479.

Wang, C., Zhang, W., Ji, Y., Marino, A., Li, C., Wang, L., Zhao, H. & Wang, M. 2024. Estimation of Aboveground Biomass for Different Forest Types Using Data from Sentinel-1, Sentinel-2, ALOS PALSAR-2, and GEDI. Forests, 15(1), 215.

Yavuz, H., Mısır, N., Mısır, M., Tüfekçioğlu, A., Karahalil, U.& Küçük, M., 2010. Karadeniz Bölgesi Saf ve Karışık Sarıçam (*Pinus slyvestris* L.) Meşcereleri için Mekanistik Büyüme Modellerinin Geliştirilmesi, Biyokütle ve Karbon Depolama Miktarlarının Belirlenmesi. TÜBİTAK TOVAG, Proje No: 106O274.

CHAPTER II

Predicting Above-ground Biomass of Pure Scots Pine Stands using Dummy Variable Regression Analysis in Northwest Türkiye

Ferhat BOLAT[1]
Sinan BULUT[2]

Introduction

Forests undertake many functions such as wood production, oxygen production, erosion prevention, landscape effects and more. Meanwhile, forests play a central role in the global carbon cycle as well as being biomass reserve areas. The biomass is a good indicator associated with the forest productivity, forest degradation, sustainable forest management strategies, carbon sequestration, and the change in forest stands over time (Puliti et al., 2021; Pietrzykowski et al., 2021). One of the most important parameters

[1] Ferhat BOLAT, Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Çankırı, Türkiye, ORCID: 0000-0003-2655-5023
[2] Sinan BULUT, Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Çankırı, Türkiye, ORCID: 0000-0001-6149-0910

that will help interpret the forest ecosystem is above-ground biomass (AGB). The AGB is the amount of living mass of trees in a forest ecosystem and provides information about the amount of currently carbon sequestration. The change in this parameter is a performance criterion for combating climate change and global warming (Bulut & Aytaş, 2023).

Recently, the efficacy of different modeling approaches in the remote sensing related studies has been frequently assessed to improve the estimation accuracy of tree and stand features, such as AGB in different forest types. Aksoy (2024) compared the prediction performance of the parametric and non-parametric models in modeling stand volume, basal area, and quadratic mean diameter of Anatolian black pine. Besides, the influence of different predictors including stand and environmental factors on the prediction accuracy has been also frequently examined, irrespective of modeling approaches. In the study of Bulut (2023), the relative contribution of different data sources such as spectral features and topographic properties in modeling AGB of Calabrian pine was investigated. It was found that an increase in the number of significant predictor variables, especially topographic and climatic factors, significantly improved the modeling accuracy.

On the other hand, stand features and site conditions (e.g., stand density classes, forest types, site classes, and ecoregion clusters) can be also included as a categorical or dummy variable in the predictive models. This approach is well examined with field data to predict tree and stand properties (e.g., total tree height and stand volume) (Manning, McDill & Gilabert, 2016; Seki & Sakıcı, 2022a; Zeng, Zhang & Tang, 2011). However, the dummy variable approach has less examined in remote sensing-related studies of forestry (Li, Li & Li, 2019; Ou et al., 2019). The present study hypothesized that determining the main source of variation in remote sensing data could provide better prediction accuracy with fewer remote sensing-based predictors. The present study aimed to analyze the influence of crown closure, stand age, and stand density classes as a dummy variable on the prediction accuracy of AGB of naturally established and managed pure Scots pine stands (*Pinus sylvestris* L.).

Material and Method

Study area

The study area is located in the Köroğlu and Dörtdivan Forest Planning Units affiliated to the Bolu Regional Directorate of Forestry, Dörtdivan Forest Enterprise in the northwest region of Türkiye (Figure 1). The geographical coordinates of this area are 401615 to 426543 east longitudes and from 4484631 to 4514094 north latitudes (WGS 84, UTM Zone 36 N). The study area covers 36433.5 ha in total, while the productive and non-productive forest areas cover 15574.8 ha and 3192.8 ha, respectively. *Pinus sylvestris* is the most common species in the area (11442.0 ha). It spreads over an area of 7569.4 ha as a pure stand, and 3872.6 ha as a mixed stand. Apart from this, *Pinus nigra* (367.2 ha) and *Quercus petraea* (129.1 ha) species are widespread. There are also areas managed as uneven aged forests with mixed stands of *Abies nordmanniana* and *Pinus sylvestris* (3636.5 ha). Elevation of the study area varies from 1160 m to 2390 m above sea level, with an average elevation of 1496 m (GDF, 2019).

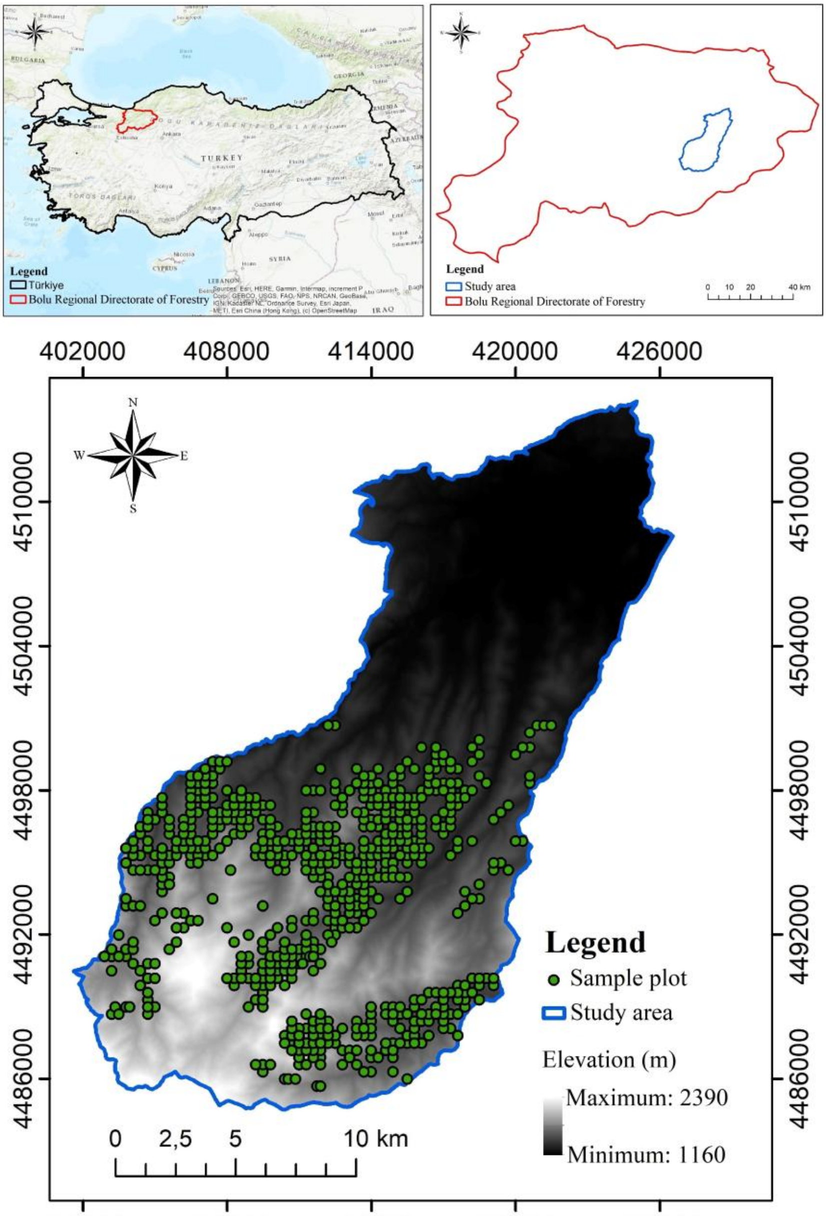

Figure 1: Location of the study area

Field Measurements

In the study, forest inventory data carried out to renew forest management plans in the forest planning units covering the study area in 2018 were used. In this context, 657 sample point data were obtained from the ground measurements. Sample plots were systematically positioned at 300 × 300 m intervals on productive forest areas. The size of sample plots was determined considering the closure classes; 800 m^2 for 11–40% closed areas, 600 m^2 for 41–70% closed areas and 400 m^2 for 71–100% closed areas. Then, all trees having diameters above 7.9 cm at breast height were identified in each sample plot and their breast height diameters (dbh, 1.30 m) were measured.

Above-ground Biomass Calculation

For the calculation of AGB in pure Scots pine stands, the diameter-dependent AGB equation developed by Yavuz et al. (2010) was used (Eq. 1). The AGB calculation for the trees with diameter measurements in each sample plot was made on a single tree basis using equation 1. Then, total AGB was determined for the sample plots. In order to prepare the data, the AGB amounts calculated as "kg" on a sample plot basis were converted to "tons" on a hectare basis.

$$AGB\ (kg) = 6.89952 - (dbh \times 2.423793) + (dbh^2 \times 0.373438) \tag{1}$$

Satellite Image Processing and Data

Landsat 8 Operational Land Imager (OLI) satellite images were downloaded free of charge from the https://earthexplorer.usgs.gov/ data portal and were used (USGS, 2000). To ensure compatibility between satellite and field data, time-series images of Landsat 8 OLI were obtained for 2018 when the inventory was made. Blue, green, red, NIR, SWIR 1, and SWIR 2 satellite bands of Landsat 8 OLI have a 30 m spatial resolution, while its panchromatic bands have a 15 m spatial resolution. Landsat 8 OLI images, which were completely cloudless, were achieved for the study area for six different months of 2018 (Table 1). Next, the Reflectance (R) data were obtained for these satellite images for each month following the atmospheric correction. Additionally, many vegetation indices (VI) were calculated using these R data. A list of the significant VIs variables in the AGB model was displayed in Table 2. Six-month R and VI data were prepared for each sample plot, and the spectral data sets were made ready for modeling.

Table 1: Description of the time-series Landsat 8 OLI data

Satellite image	Path/Row	Acquisition data
Landsat 8 OLI	178/032	17 March 2018 28 June 2018 14 July 2018 15 August 2018 16 September 2018 3 November 2018

Table 2: Vegetation indices derived from Landsat 8 OLI satellite image

Vegetation indices	Formula	Reference
EVI (Enhanced Vegetation Index)	2.5 x ((NIR – Red) / (NIR + (6 x Red) – (7.5 x Blue) + 1))	Liu & Huete (1995)
SARVI (Soil and Atmospherically Resistant Vegetation Index)	(1.0 + L) x (NIR - (Rr – y) x (RB - Rr)) / (NIR - ((Rr – y) x (RB - Rr)) + L) *y = 0.735, Rr = 0.740, L = 0.487, RB = 0.560*	Kaufman & Tanre (1992)
VIS123	(Blue + Green + Red)	Lu et al. (2004)

Dummy Variable Regression Analysis

Prior to developing the predictive AGB model, a suitable set of reflectance variables and vegetation indices was determined using the forward variable selection method on the basis of Akaike's information criterion. At this time, the collinearity between the potential predictors was tested using the variance inflation factor (VIF) with a threshold of 10. After determining the significant predictors, the present study utilized dummy variable approach to predict AGB of Scots pine stands. This method has an ability to involve to the factorized effect of different sites and stand features such as ecoregions and crown closure classes (Corral-Rivas et al., 2007; Seki & Sakıcı, 2022a, 2022b). Therefore, it significantly improves modeling accuracy, as compared to traditional regression models. The general form of dummy variable models is as follows (Eq. 2):

$$Y = a_0 + \sum a_i z_i + bX \quad (2)$$

where Y is a dependent variable, X is a predictor variable, a and b are regression parameters to be estimated, a_i is a specific parameter to be estimated, and z_i is indicator variable tailored to a categoric factor such as crown closure classes.

The indicator variables that have the values 0 and 1 are usually used to define dummy variables. Several schemes can be used to assign dummy variable (Zeng, Zhang & Tang, 2011). In the current study, dummy variable approach with reference group was performed in developing AGB models. In pre-analysis, among different classes including crown closure (CC), stand age (SA) and stand density (SD), the best AGB predictions were achieved by SD classes that were represented by the categorical variable with three categories using two dummy variables. Stand density of sample plots was calculated using an equation developed for Scots pine stands by Yavuz et al. (2010). Accordingly, sample plots were divided into three categories involving SD <0.3—SD_1, 0.3<SD<0.6—SD_2 and SD>0.6—SD_3.

Eq. 2 was then extended with dummy variables to include SD classes. The resultant form of AGB model is as follows (Eq. 3):

$$Y = a_0 + a_1 I_1 + a_2 I_2 + (b_0 + b_1 I_1 + b_2 I_2)X \tag{3}$$

where I_i corresponds dummy variables as follows:

I_1=1 and I_2=0 for SD_1

I_1=0 and I_2=1 for SD_2

I_1=0 and I_2=0 for SD_3 (reference group)

Model Performance Criteria

The present study utilized the coefficient of determination (R^2, Eq. 4) and mean absolute percentage error (MAPE, Eq. 5) to assess modeling accuracy. After the best models were determined based on these criteria, their predictive ability was then justified using the test data.

$$R^2 = 1 - \frac{\sum(AGB - \widehat{AGB})^2}{\sum(AGB - \overline{AGB})^2} \tag{4}$$

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{AGB - \widehat{AGB}}{\overline{AGB}} \right| \tag{5}$$

where the hat corresponds with the predicted values and the overline shows the mean of actual values.

Results

The results of dummy variable regression models associated with various categories were shown in Table 3. The best accuracy was gained by SD as a dummy variable. Therefore, in the following pages, the results based on only SD were provided.

Table 3: The performance scores for AGB models with respect to different categories in model data

Data source	Categories	R^2	MAPE
Reflectance	CC	0.18	26.16
	SA	0.32	24.22
	SD	0.74	15.23
Vegetation indices	CC	0.15	27.32
	SA	0.36	23.59
	SD	0.73	15.33

CC: Crown closure, SA: Stand age, SD: Stand density

The results of regression model with and without dummy variable were displayed in Table 4. As seen in the table, the inclusion of dummy variable into regression model considerably improved modeling accuracy. In addition, the success of this model in test data was found to be acceptable.

Table 4: The performance scores of AGB model without (M1) and with dummy variable (M2) in model and test data in terms of reflectance variables

Performance Criteria	M1		M2	
	Model Data	Test Data	Model Data	Test Data
R^2	0.14	0.10	0.74	0.67
MAPE	27.13	27.90	15.23	17.74

The model forms involving reflectance variables of linear regression (Eq. 6) and dummy variable regression (Eq. 7) were presented below.

- Regression model with reflectance variables (M1)

$$\text{AGB ton ha}^{-1} = (a_1 \times \text{SWIR } 2_{\text{July}}) + (a_2 \times \text{NIR}_{\text{March}}) + (a_3 \times \text{Blue}_{\text{March}}) \quad (6)$$

- Dummy variable regression model with reflectance variables (M2)

$$\text{AGB ton ha}^{-1} = (a_1+a_2\times I_1+a_3\times I_2) + (a_4+a_5\times I_1+a_6\times I_2) \times \text{SWIR2}_{\text{July}} + (a_7+a_8\times I_1+a_9\times I_2) \times \text{NIR}_{\text{March}} + (a_{10}+a_{11}\times I_1+a_{12}\times I_2) \times \text{Blue}_{\text{March}} \quad (7)$$

Similar to the previous results, utilizing dummy variable regression model with vegetation indices significantly increased the modeling accuracy (Table 5). In addition, the successes of the reflectance-based models were close to each other. The same result was valid to the vegetation indices-based models.

Table 5: The performance scores of AGB model without (M3) and with dummy variable (M4) in model and test data in terms of vegetation indices variables

Performance Criteria	M3		M4	
	Model Data	Test Data	Model Data	Test Data
R^2	0.15	0.14	0.73	0.65
MAPE	27.33	27.66	15.33	17.91

The model forms involving vegetation indices of linear regression (Eq. 8) and dummy variable regression (Eq. 9) were presented below.

- Regression model with vegetation indices (M3)

$$AGB\ ton\ ha^{-1} = (a_1 \times EVI_{July}) + (a_2 \times VIS123_{June}) + (a_3 \times SARVI_{August}) \quad (8)$$

- Dummy variable regression model with vegetation indices (M4)

$$AGB\ ton\ ha^{-1} = (a_1 + a_2 \times I_1 + a_3 \times I_2) + (a_4 + a_5 \times I_1 + a_6 \times I_2) \times EVI_{July} + (a_7 + a_8 \times I_1 + a_9 \times I_2) \times VIS123_{June} + (a_{10} + a_{11} \times I_1 + a_{12} \times I_2) \times SARVI_{August} \quad (9)$$

The estimated parameters of dummy variable regression model involving reflectance (M2) and vegetation indices (M4) were displayed in Table 6. As shown in the table, the coefficients belonging to SD classes were very different, suggesting the presence of a considerable variation among SD classes in terms of reflectance and vegetation indices.

Table 6: The estimated parameters of dummy variable regression models including reflectance (M2) and vegetation indices (M4) variables as predictor variables

Parameters	Coefficients		Parameters	Coefficients	
	M2	M4		M2	M4
a_1	215.76	171.37	a_7	-222.40	-76.27
a_2	-165.51	-156.92	a_8	235.67	187.65
a_3	-70.25	-26.15	a_9	159.03	16.83
a_4	28.11	-107.52	a_{10}	294.05	11.19
a_5	9.15	84.18	a_{11}	-292.78	-6.34
a_6	-103.29	19.59	a_{12}	-265.45	-9.05

The mathematical expressions of dummy variable regression models, that is M2 and M4, were given below to clarify their use.

A) Dummy variable regression models having reflectance variables were as follows (Eq. 10-12):

For SD_1:

$$AGB\ ton\ ha^{-1} = 50.25 + (37.26 \times SWIR2_{July}) + (13.27 \times NIR_{March}) + (1.27 \times Blue_{March}) \quad (10)$$

For SD_2:

$$AGB\ ton\ ha^{-1} = 145.51 - (75.18 \times SWIR2_{July}) - (63.37 \times NIR_{March}) + (28.16 \times Blue_{March}) \quad (11)$$

For SD_3:

$$AGB\ ton\ ha^{-1} = 215.76 + (28.11 \times SWIR2_{July}) - (222.40 \times NIR_{March}) + (294.05 \times Blue_{March}) \quad (12)$$

B) Dummy variable regression model having vegetation indices were as follows (Eq. 13-15):

For SD_1:

$$AGB\ ton\ ha^{-1} = 20.45 - (23.34 \times EVI_{July}) + (111.38 \times VIS123_{June}) + (4.85 \times SARVI_{August}) \quad (13)$$

For SD_2:

$$AGB\ ton\ ha^{-1} = 145.22 - (87.93 \times EVI_{July}) - (59.44 \times VIS123_{June}) + (2.14 \times SARVI_{August}) \quad (14)$$

For SD_3:

$$AGB\ ton\ ha^{-1} = 171.37 + (28.11 \times EVI_{July}) - (222.40 \times VIS123_{June}) + (11.19 \times SARVI_{August}) \quad (15)$$

Discussion

The present study found that the R^2 and MAPE values of reflectance-based models were the closest to that of vegetation-based models, which were approximately 0.15 and 27% for linear regression and 0.73 and %15 for dummy variable regression, respectively. The improved modeling accuracy with the dummy variables can be attributed to the different structural properties of forests with respect to canopy, density (overstory and understory), and species composition. Furthermore, it can be related to the different ecological characteristics of forests regarding climate and topography. The varying stand structures and growing conditions greatly affect the spectral characteristics of forest stands. The stand features such as stand volume and AGB are thus reflected by different colors, structures and textures in the satellite images (Li, Li & Li, 2019). Andalibi et al. (2021) found that the vegetation indices including EVI and NDVI had great spatial and temporal variations in a semi-arid forest. In another study, Hossain & Li (2021) found that NDVI significantly varied between the study areas including the temperate forests located in the cold steppe, humid and dry steppe ecoregions. The high reflectance variability presents a significant challenge to the traditional linear regression models and leads to the low prediction accuracy, as evidenced by the current study. In this circumstance, the non-parametric models such as the support vector machine have been preferred from the linear regression model. However, they could explain only a small proportion of the variation in stand features when available data is insufficient (Ou et al., 2019). Therefore, considering the effect of forests structural properties as a dummy variable can improve modeling accuracy. Ou et al. (2019) found that the inclusion of dummy variables relating to stand age in linear regression increased the R^2 value from 0.32 to 0.70. Likewise, Li, Li & Li (2019) found that incorporating the dummy variables relating to the crown density into linear model increased the R^2 value from 0.20 to 0.40 in the pine and mixed forests, and 0.20 to 0.60 in the Chinese fir forests. The current study underlined that the stand density was the main source of the variation in remote-sensing data reflecting AGB of Scots pine stands. When the stand density classes were included as a dummy variable in the predictive AGB model, the modeling accuracy greatly improved. The dummy variable regression model explained a large proportion of the variation in the observed AGB, as the R^2 was higher than 70%.

Satellite images-based R and VI data were widely used for predicting forest AGB. Bulut & Aytaş (2023) predicted the AGB of Scots pine, which is distributed in the Inner Anatolian Region of Türkiye, using different data sets. NIR, MIR and EVI were added as the auxiliary predictor variables to the linear regression model using spectral data obtained from the MODIS satellite (R^2_{adj}=0.46, r=0.69, RMSE=18.002 t ha^{-1}). In addition, the highest improvement in the AGB model was achieved by using spectral variables. Bulut (2023) predicted the AGB of Calabrian pine stands spreading in the Mediterranean area of Türkiye. The R and VI data sets were used as the auxiliary predictor variables in the multiple linear regression model (MLR) and support vector machines (SVM) techniques. In these AGB models, the bands included as independent variables were blue, NIR, SWIR-1 and SWIR-2, and the indices were NVWI, PSSR, SARVI, GLI, ARVI, EVI and GNDVI. The R^2 values were obtained as 0.34 and 0.50 using the MLR method with the R and VI data sets. When the SVM method was used, these values were 0.54 and 0.61, respectively.

Turgut & Günlü (2022) predicted the AGB of Anatolian black pine stands spreading in the north of Türkiye. The AGB was predicted using the MLR method with band brightness (R^2_{adj}=0.445), VI (R^2_{adj}=0.387) and texture (R^2_{adj}=0.552) values derived from Landsat 8 OLI satellite. In the AGB model, the bands included as independent variables were blue, red, SWIR-1 and TIR-2, and the indices were FII, DVI, EVI and IPVI. In this study, although the prediction performance of R data is higher than VI data, VI variables can be used effectively in AGB prediction. In the study conducted by Bulut (2023), it was stated that the addition of Landsat 8 OLI satellite based VI variables to the model provided the most contribution to the AGB model. VI can minimize soil background, sun angles, topography effects, atmospheric variability, and canopy geometry. It can also significantly increase the sensitivity to green vegetation. This situation might improve the correlation between VI and AGB in forest ecosystems. So, VI can be used as an alternative variable to predict the AGB (Günlü et al., 2014; Nguyen, Vu & Roeder, 2021).

In modeling using remotely sensed data, the use of time-series data instead of single-dated data is very important. Zhu & Liu (2015) reported that phenological features reflected by time-series satellite data instead of single-dated data can reduce saturation problems and models with highly predictive can be developed. According to Chrysafis et al. (2019), time-series spectral data show phenological variations across a range of vegetation species. Additionally, they found that time-series spectral data outperform single-dated data in estimating growing stock volume. Spectral data is influenced by forest phenology, and as a result, spectral variations that may exist in various forest types may have an impact on the prediction accuracy of the models that need to be developed. In these situations, predicting forest stand parameters may not be possible with just single-time spectral data. Utilizing spectral data from satellite imagery in time-series helps lessen saturation and improve the accuracy of models that are being created for various vegetation structures (Naik, Dalponte & Bruzzone, 2021; Bulut, Sivrikaya & Günlü, 2022).

Conclusion

The current study drew attention to the role of dummy variable regression model in order to improve the AGB predictions of Scots pine stands. The classification of the AGB and spectral data by a certain stand feature such as crown closure, stand age, and stand density significantly improved the prediction accuracy. The present study showed that the main factor as a dummy variable was the stand density. When the influence of stand density was included as a dummy variable in the predictive AGB model, the accuracy of the AGB predictions significantly improved, as the R^2 increased from 0.15 to 0.70 and the MAPE reduced from 27.3 to 15.3. It is suggested utilizing the dummy variable regression with an appropriate categorical factor in order to improve the quality of AGB predictions.

Acknowledgements

The authors thanks to the General Directorate of Forestry of Turkey for providing data.

References

Aksoy, H. 2024. Estimation Stand Volume, Basal Area and Quadratic Mean Diameter Using Landsat 8 OLI and Sentinel-2 Satellite Image with Different Machine Learning Techniques. Transactions in GIS, n/a(n/a), 1–18. https://doi.org/10.1111/tgis.13265

Andalibi, L., Ghorbani, A., Moameri, M., Hazbavi, Z., Nothdurft, A., Jafari, R., & Dadjou, F. (2021). Leaf Area Index Variations in Ecoregions of Ardabil Province, Iran. Remote Sensing, 13(15), 2879. https://doi.org/10.3390/rs13152879

Bulut, S. 2023. Machine learning prediction of above-ground biomass in pure Calabrian pine (Pinus brutia Ten.) stands of the Mediterranean region, Türkiye. Ecological Informatics, 74, 101951.

Bulut, S. & Aytaş, İ. 2023. Modeling potential distribution and above-ground biomass of Scots pine (Pinus sylvestris L.) forests in the Inner Anatolian Region, Türkiye. Environmental Monitoring and Assessment, 195(12), 1471.

Bulut, S., Sivrikaya, F. & Günlü, A. 2022. Evaluating statistical and combine method to predict stand above-ground biomass using remotely sensed data. Arab. J. Geosci. 15, 838.

Chrysafis, I., Mallinis, G., Tsakiri, M. & Patias, P. 2019. Evaluation of single-date and multi-seasonal spatial and spectral information of Sentinel-2 imagery to assess growing stock volume of a Mediterranean forest. Int. J. Appl. Earth Obs. Geoinf. 77, 1–14.

Corral-Rivas, J. J., Barrio-Anta, M., Aguirre-Calderón, O. A. & Diéguez-Aranda, U. 2007. Use of stump diameter to estimate diameter at breast height and tree volume for major pine species in

El Salto, Durango (Mexico). Forestry: An International Journal of Forest Research, 80(1), 29–40. https://doi.org/10.1093/forestry/cpl048

GDF, 2019. Dörtdivan and Köroğlu Forest Planning Units, Forest Management Plan. General Directorate of Forestry, Ankara.

Günlü, A., Ercanli, I., Başkent, E.Z. & Çakır, G. 2014. Estimating aboveground biomass using Landsat TM imagery: a case study of Anatolian Crimean pine forests in Turkey. Ann. Forest Res. 57, 289–298.

Hossain, M. L. & Li, J. 2021. NDVI-based vegetation dynamics and its resistance and resilience to different intensities of climatic events. Global Ecology and Conservation, 30, e01768. https://doi.org/10.1016/j.gecco.2021.e01768

Kaufman, Y.J. & Tanre, D. 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. IEEE transactions on Geoscience and Remote Sensing, 30(2), 261-270.

Li, C., Li, Y. & Li, M. 2019. Improving Forest Aboveground Biomass (AGB) Estimation by Incorporating Crown Density and Using Landsat 8 OLI Images of a Subtropical Forest in Western Hunan in Central China. Forests, 10(2), 104. https://doi.org/10.3390/f10020104

Liu, H.Q. & Huete, A. 1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE transactions on geoscience and remote sensing, 33(2), 457-465.

Lu, D., Mausel, P., Brondizio, E. & Moran, E. 2004. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. Forest Ecology and Management, 198(1–3): 149–167.

Manning, P. J., McDill, M. E. & Gilabert, H. 2016. An alternative model for describing even-aged stand-level sawtimber growth in Pennsylvania. Forest Science, 62(6), 663–670. https://doi.org/10.5849/forsci.15-163

Naik, P., Dalponte, M. & Bruzzone, L. 2021. Prediction of Forest aboveground biomass using multitemporal multispectral remote sensing data. Remote Sens. 13, 1282.

Nguyen, H.H., Vu, H.D. & Roeder, A. 2021. Estimation of above-ground mangrove biomass using Landsat-8 data- derived vegetation indices: a case study in Quang Ninh Province, Vietnam. For. Soc. 5, 506–525.

Ou, G., Li, C., Lv, Y., Wei, A., Xiong, H., Xu, H. & Wang, G. 2019. Improving Aboveground Biomass Estimation of Pinus densata Forests in Yunnan Using Landsat 8 Imagery by Incorporating Age Dummy Variable and Method Comparison. Remote Sensing, 11(7), 738. https://doi.org/10.3390/rs11070738

Pietrzykowski, M., Wo´s, B., Tylek, P., Kwa´sniewski, D., Juliszewski, T., Walczyk, J., Likus-Cie´slik, J., Ochał, W. & Tabor, S. 2021. Carbon sink potential and allocation in above- and below-ground biomass in willow coppice. J. For. Res. 32, 349–354.

Puliti, S., Breidenbach, J., Schumacher, J., Hauglin, M., Klingenberg, T.F. & Astrup, R. 2021. Above-ground biomass change estimation using national forest inventory data with Sentinel-2 and Landsat. Remote Sens. Environ. 265, 112644.

Seki, M. & Sakıcı, O.E. 2022a. Ecoregional variation of Crimean pine (Pinus nigra subspecies pallasiana [Lamb.] Holmboe) stand growth. Forest Science, 68(5–6), 452–463. https://doi.org/10.1093/forsci/fxac030

Seki, M. & Sakıcı, O.E. 2022b. Ecoregion-based height-diameter models for Crimean pine. Journal of Forest Research, 27(1), 36–44. https://doi.org/10.1080/13416979.2021.1972511

Turgut, R. & Günlü, A. 2022. Estimating aboveground biomass using Landsat 8 OLI satellite image in pure Crimean pine (Pinus nigra J.F. Arnold subsp. Pallasiana (lamb.) Holmboe) stands: a case from Turkey. Geocarto Int. 37, 720–734.

USGS, 2000. USGS Earth Explorer. URL. https://earthexplorer.usgs.gov/ (accessed 3.07.24).

Yavuz, H., Mısır, N., Mısır, M., Tüfekçioğlu, A., Karahalil, U. & Küçük, M. 2010. Karadeniz Bölgesi saf ve karışık sarıçam (Pinus Slyvestris L.) meşcereleri için mekanistik büyüme modellerinin geliştirilmesi, biyokütle ve karbon depolama miktarlarının belirlenmesi (1001-Araştırma Projesi No. 106O274) (p. 318). TÜBİTAK.

Zeng, W. S., Zhang, H. R. & Tang, S.Z. 2011. Using the dummy variable model approach to construct compatible single-tree biomass equations at different scales — a case study for Masson pine (Pinus massoniana) in southern China. Canadian Journal of Forest Research, 41(7), 1547–1554.

Zhu, X. & Liu, D. 2015. Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. ISPRS J. Photogramm. Remote Sens. 102, 222–231.

CHAPTER III

Real Time Detection of Forest Fires with YOLO Algorithm Using Digital Forest Images: A Deep Learning-Based Approach

Hasan AKSOY[1]
Enes CENGİZ[2]

Introduction

Forests are biologically rich ecosystems (Zhou et al., 2022), that provide wood, non-wood products, food, biofuels, and pharmaceutical raw materials for more than one billion people (FAO, 2020; Nesha et al., 2021), host wildlife (Gibson et al., 2011; Sasaki, 2021) and different plant species (Sullivan et al., 2017; Sasaki, 2021). Beyond these services, it is essential for sequestering carbon, contributing significantly to the mitigation of global climate change (Pan et al., 2011; Nesha et al., 2021). Therefore, monitoring

[1] Hasan AKSOY, Sinop University, Ayancık Vocational School, Department of Forestry, Forestry and Forest Products Program, Sinop, Türkiye. haksoy@sinop.edu.tr. ORCID: 0000-0003-1980-3834

[2] Enes CENGİZ, Sinop University, Ayancık Vocational School, Department of Electronics and Automation, Control and Automation Technology Program, Sinop, Türkiye. ecengiz@sinop.edu.tr. ORCID: 0000-0003-1127-2194

and controlling forest structure is very important for the sustainability of forests (Wang et al., 2018; Aksoy and Kaptan, 2022). Fires have historically played the greatest role in changing forest ecosystems and jeopardizing their continuity and sustainability (Bond et al., 2005; Glasspool et al., 2004: Sivrikaya et al., 2024). In recent times, driven by global climate change, the escalation in the occurrence and intensity of forest fires has become a critical threat to both forest ecosystems and society. This is because these fires not only lead to the degradation of forests and other vegetation and the loss of habitats, but can also have a profound impact on the ecological balance and the global climate (Thonicke et al., 2001; Yadav et al., 2022; Yang et al., 2024). Extended periods of hot, dry summers greatly elevate the risk of forest fires, particularly in coniferous pine forests and maquis regions, which are highly susceptible to large-scale blazes (Bilgili et al., 2021; Sivrikaya et al., 2024). As a result, forest fires have become a major problem in recent years, particularly in the countries of the Mediterranean basin and in Turkey. The development of fire monitoring technologies for the early detection and swift response to forest fires has become an essential and urgent necessity. (Yang et al., 2024).

The methods commonly used for forest fire detection can be summarized under three main headings: forest patrols, satellite systems, and video surveillance (Xie et al., 2018; Barmpoutis et al., 2020; Cao et al., 2024). Ground patrols are conducted on foot or by vehicle and can only be carried out in limited areas. This limits visibility and is insufficient for fire detection. Satellite systems, on the other hand, eliminate the disadvantage of limited space, but provide a long, costly monitoring time and a low-resolution (coarse texture) detection service. In addition, satellite systems can be affected by weather and cloud cover, making them inadequate for real-time fire detection (Chuvieco et al., 2020; Yang et al., 2024). Recently, researchers have focused on investigating the real-time detection of fires using in-forest cameras (photo traps) that minimize the aforementioned limitations and unmanned aerial vehicles (UAVs), which have recently developed rapidly and are widely used in many fields (Chen et al., 2019; Mahmoud et al., 2019; Yang et al., 2024). Mukhiddinov et al. (2022) investigated the early detection of forest fires based on smoke detection in images using UAV imagery. Emmy Prema et al. (2018) investigated fire detection by examining both static and dynamic texture features in areas recognized as fire zones, employing the YVbCr color model. Moreover, fire detection using RGB, HIS and YUV multicolor sensors has been investigated by Han et al. (2017).

The diverse range of data made available through advancements in remote sensing has motivated researchers to incorporate widely used machine learning and deep learning algorithms into remote sensing methodologies (Özer et al., 2022; Aysal et al., 2022). The key algorithms for detecting forest fires have been developed using deep learning methods (LeCun et al., 2015; Cao et al., 2024). Among these is the region-based convolutional neural network (R-CNN), a two-stage method known for its high accuracy in identifying regional targets, albeit with a slower processing speed (Girshick et al., 2014; Ren, 2015; Zou et al., 2023). Others are regression-based single-stage algorithms such as Center-Net, SSD, R-SSD and You Only Look Once (YOLO) series, which are fast but slightly less accurate (Redmon et al., 2016; Liu et al., 2016; Jeong et al., 2017; Song and Fu, 2018). Two-stage and one-stage approaches have been applied in forest fire detection (Tan et al., 2020; Cheknane et al., 2024; Cao et al., 2024). However, the YOLO algorithm is the most common algorithm used in forest fire detection as a CNN-based object detector commonly used in computer vision applications. Lin et al. (2023) proposed TCA-YOLO, a highly efficient and precise model for global forest fire detection. This model integrates YOLOv5 with a Transformer encoder to enhance detection capabilities. TCA-YOLO exhibited very high predictive ability in wildfire detection in various scenarios. Li et al., (2023) demonstrated approaches to facilitate forest fire management by improving forest fire and smoke detection. Abdusalomov et al. (2023) explored the use of the Detectron2 model and deep learning techniques for detecting forest fires. Similarly, Lu et al. (2022) examined real-time fire detection employing UAVs and deep learning methods. The YOLOv8 model, as discussed by Talaat and ZainEldin (2023), has emerged as an effective solution for real-time forest fire detection, offering notable advancements in both accuracy and speed. This study aims to detect forest fires early by using the YOLO algorithm with mild, medium, and heavy digital fire images of forest fires obtained from UAVs, photo traps, and other camera systems. Thus, forest fires can be detected early and in real time with sensors such as UAVs and photo traps, which are being used extensively, thus contributing significantly to the first response time to forest fires.

Material and Method

This research focuses on the real-time, high-precision detection of forest fires, accomplished through three primary phases. The initial phase entails the development of fire image datasets, followed by the second phase, which focuses on model training, and culminating in the final phase, which evaluates the accuracy of model. The general flow of the study methodology is shown in Figure 1.

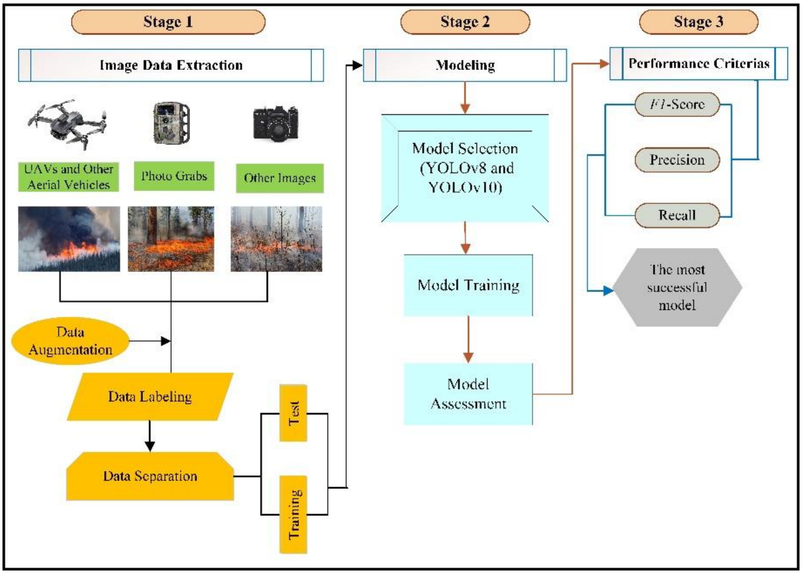

Figure 1. Flowchart representing the process of the study

Data Set and Image Processing

The first stage of the overall study methodology is data acquisition. At this stage, light, medium, and high intensity fire images obtained from different sensors such as drones, photo traps, cameras, etc. were obtained from the Kaggle public dataset platform. Some of the data was also gathered from publicly available sources, including social media and news websites. The obtained images were carefully edited and checked for repetition to avoid using the same images. All images were manually annotated based on whether or not fires were present (Cengiz et al., 2022). In total, 1020 images were used for the analysis. Finally, 80% (816 images) of the total image dataset was used for training and 20% (204 images) for testing (Yıldız and Serttaş, 2023; Kelek et al., 2021). Some of the fire images used in the modeling are shown in Figure 2.

Figure 2. Fire image series used in modeling

Deep Learning (YOLO Algorithm)

Choosing the right object detection algorithm is essential for developing an effective fire detection model. Options such as YOLOv8, Faster R-CNN, and SSD each come with their own set of benefits and drawbacks. The selected algorithm should align well with the dataset and be proficient in identifying different fire scenarios to fulfill the requirements of the intelligent fire detection system. YOLOv8 and YOLOv10 are the most preferred algorithms in fire detection systems due to their speed and accuracy (Talaat and ZainEldin, 2023). In this study, these two algorithms are used for the forest fire detection system. The YOLOv8 and YOLOv10 models were trained on the prepared labeled dataset. Model training consists of instructing the deep learning model to identify the characteristics of fire images and accurately distinguish them. The methodology of the YOLOv8 model is shown in Figure 3 and the methodology of the YOLOv10 model is shown in Figure 4. There are three basic versions of YOLO algorithms: s (small), m (medium), and x (extra-large).

YOLO-s (small): The smallest model and ideal for achieving fast results on lightweight hardware. This model is preferred in applications that require high speed and low memory consumption at the expense of accuracy.

YOLO-m (medium): Offers a balanced structure. Suitable for users looking for a balance between performance and accuracy. Increases accuracy with more parameters while remaining efficient in terms of speed.

YOLO-x (extra-large): This is the larger model compared to the other two versions and is designed to provide the highest performance. It is especially suitable for use in systems that require high performance and have more computing power.

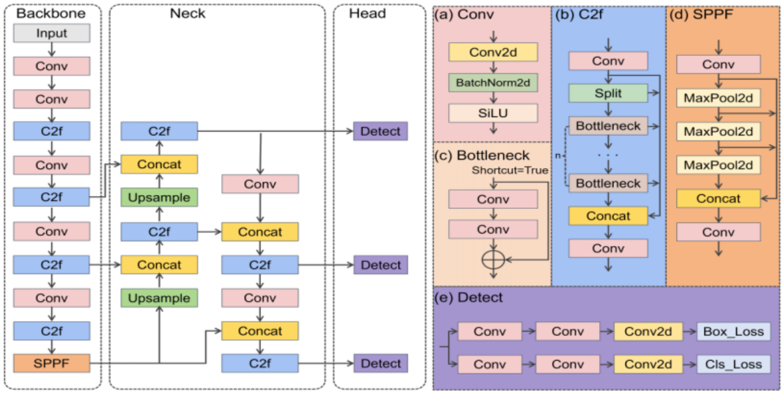

Figure 3. YOLOv8 algorithm working methodology (Talaat and ZainEldin, 2023)

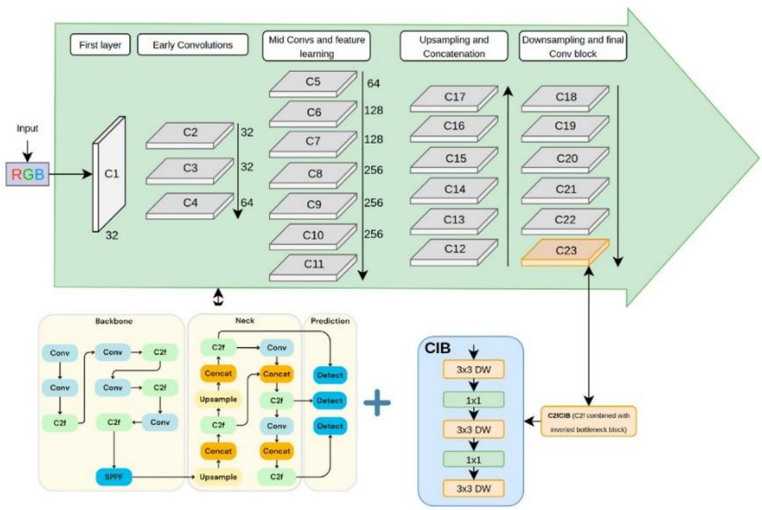

Figure 4. YOLOv10 algorithm working methodology (Akhmedov et al., 2024)

Evaluation of Fire Detection Performance of Models

The complexity matrix is used to examine the performance criteria of YOLO models in the classification process. The complexity matrix is given in Figure 5. The fire detection performance levels of the images obtained as a result of modeling were checked with precision, sensitivity, and F1-score. Precision is the ratio of correctly classified images to the sum of misclassified and negatively classified images. Sensitivity is defined as the proportion of accurately classified images to the total number of images. Recall is also known as sensitivity or specificity. F1-score is an evaluation criterion expressed as the harmonic mean of precision and recall (Yacouby and Axman, 2020). The mathematical equivalents of precision, recall, and F1-score are given in Equation (1-3) respectively.

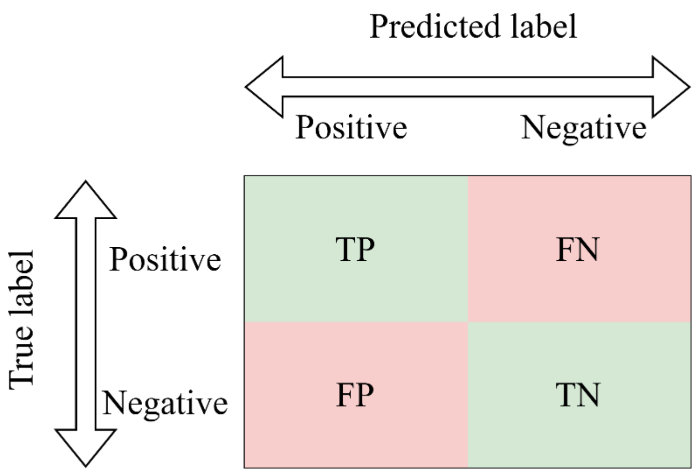

Figure 5. Confusion matrix

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{F1} - \text{score} = 2 \text{ x } \frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

Result

The experimental outcomes of the study are presented in this section. YOLOv8 and YOLOv10 models were given 80% of the images in the dataset for training. Then the models were tested with 20% of the remaining dataset. The complexity matrix of 6 different models (YOLOv8-s, YOLOv8-m, YOLOv8-x, YOLOv10-s, YOLOv10-m, and YOLOv10-x) are given in Figure 6.

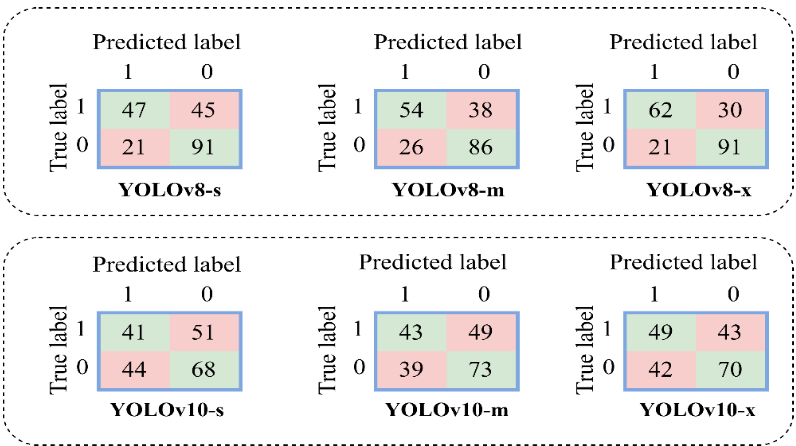

Figure 6. Confusion matrix results of YOLOv8 and YOLOv10 models

The performance criteria in Equation (1-3) were calculated by taking the TP, FN, FP, and TN values in the confusion matrix given in Figure 6. The precision, recall and F1-score values of the models according to the calculation results are presented in Table 1.

Table 1. Presentation of performance parameters of YOLO models

Model	Precision	Recall	F1-score
YOLOv8-s	0.69	0.51	0.59
YOLOv8-m	0.68	0.59	0.63
YOLOv8-x	0.75	0.67	0.71
YOLOv10-s	0.48	0.45	0.46
YOLOv10-m	0.52	0.47	0.49
YOLOv10-x	0.54	0.53	0.54

The results in Table 1 compare the forest fire detection performance of YOLOv8 and YOLOv10 algorithms in terms of precision, recall, and F1-score parameters. YOLOv8 models generally showed higher performance. In particular, the YOLOv8-x model achieved the highest performance with 75% precision, 67% recall, and 71% F1-score. Since this model has a larger structure than the other YOLOv8 variants, it is the most suitable model for applications requiring high accuracy. On the other hand, YOLOv10 models showed lower performance than YOLOv8. The YOLOv10-s model has the lowest precision (48%) and F1-score (46%) and can be used in applications where high speed is required but accuracy is relatively less important. In general, YOLOv8 models appear to be more balanced and accurate, making them a better choice for early and accurate detection of forest fires. The outcomes of the fire detection analysis in the study are illustrated in Figure 7.

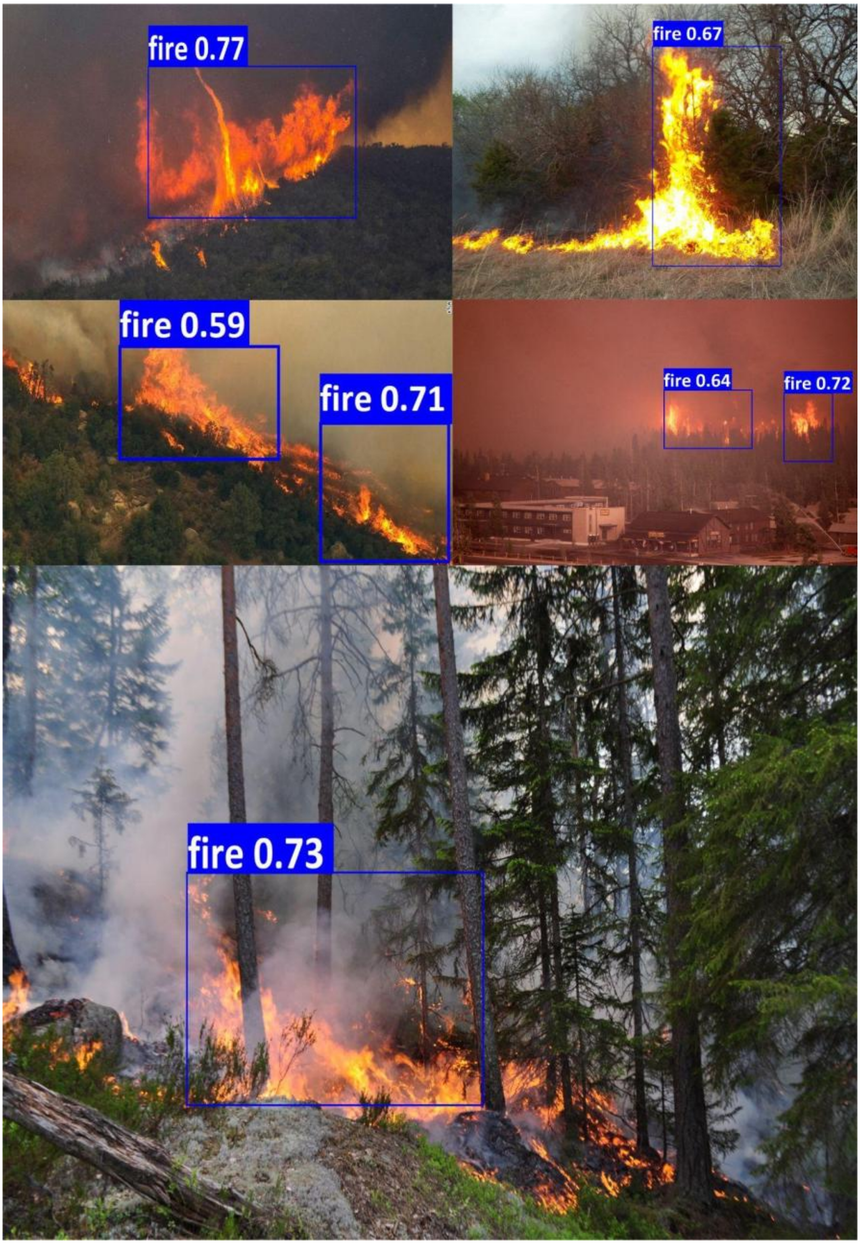

Figure 7. The result images of the models used for fire detection

Discussion

This study compares two versions of the YOLO algorithm, YOLOv8, and YOLOv10, for forest fire detection. The findings of the study show that YOLOv8-x gives the most successful results with high precision, recall, and F1-score values compared to the other versions. These findings support the achievements of Talaat and ZainEldin (2023) in fire detection using YOLOv8. The high accuracy and speed of YOLOv8 make it a viable option for early detection systems for incidents that require immediate response such as forest fires.

The results of the study are in line with the existing literature. For example, Lin et al. (2023) reported that the TCA-YOLO model, which they created by combining YOLOv5 and a Transformer encoder, offers high accuracy in fire detection. The accuracy advantages of YOLO-based models are highlighted by their fast response times compared to other single-stage algorithms (SSD, CenterNet) (Liu et al., 2016; Song and Fu, 2018). Fast and accurate analysis, especially in digital forest fire images obtained from UAV and camera systems, can support early intervention and minimize forest losses.

This study is important in terms of emphasizing the performance difference between versions of YOLO algorithms. In previous studies, Mukhiddinov et al. (2022) used the YOLOv5 model for smoke detection on UAV images and obtained successful results in detecting fires at an early stage. However, in this study with YOLOv8 and YOLOv10 models, it was revealed that YOLOv8 is superior in terms of accuracy and sensitivity. In particular, the fact that YOLOv8-x was trained with a wider set of parameters increased the accuracy of the model in fire detection. This is a critical advantage, especially for systems that can analyze at high resolution and process more data.

However, the study has some limitations. Although YOLOv10 models provide a speed advantage, they are characterized by low accuracy rates. This limits its use in emergency response systems that require high accuracy. Abdusalomov et al. (2023) obtained successful results in fire detection using the Detectron2 model but suggested that faster algorithms such as YOLO are preferable for applications with higher speed requirements. Accordingly, the low accuracy of YOLOv10 may be suitable for scenarios where speed requirements are higher but accuracy is less important.

In conclusion, the findings of the study show that the YOLOv8-x model is one of the best options for forest fire detection. With its high precision and recall, this model provides a reliable solution for fire detection applications. In future studies, the improved versions of YOLOv8 and YOLOv10 can be examined in more complex fire detection scenarios to provide suggestions for performance improvement. In addition, testing with high-resolution satellite imagery and multiple data sources could increase the sensitivity and overall accuracy of the models to different types of data.

Conclusion

In this study, a dataset composed of digital images obtained from UAVs, camera traps, and various imaging devices was utilized to employ deep learning models for the early detection of forest fires. Training datasets were applied on three different versions (s, m, and x) of the deep learning models YOLOv8 and YOLOv10. The models were then evaluated using a test dataset, and the results were presented in a confusion matrix. The experimental results showed that the YOLOv8 model outperformed the YOLOv10 model. In particular, the YOLOv8-x model (F1-score, 0.71), with high precision and recall rates, emerged as a reliable option for fire detection. These results demonstrate the viability of deep learning-based fire detection systems in enabling rapid response capabilities in emergency management and forest fire control.

In future studies, further improvements could be made to achieve higher accuracy and adaptability in this field. Firstly, advanced versions of the YOLO algorithms or comparisons with other deep learning models could be considered. In particular, the future versions of YOLOv8 and YOLOv10 could be tested on more complex datasets and high-resolution images to enhance their performance. Additionally, data diversity could be increased; for example, integrating information on smoke propagation and wind speed along with fire area imagery could lead to a more comprehensive fire detection system. The combination of multiple data sources, such as drone and satellite imagery, camera traps, and meteorological data, would allow the model to better adapt to various scenarios. This would result in more accurate outcomes in determining fire risk.

References

Abdusalomov, A. B., Islam, B. M. S., Nasimov, R., Mukhiddinov, M. & Whangbo, T. K. (2023). An improved forest fire detection method based on the detectron2 model and a deep learning approach. *Sensors, 23*(3), 1512.

Akhmedov, F., Nasimov, R. & Abdusalomov, A. (2024). Dehazing Algorithm Integration with YOLO-v10 for Ship Fire Detection. *Fire, 7*(9), 332.

Aksoy, H. & Kaptan, S. (2022). Simulation of future forest and land use/cover changes (2019–2039) using the cellular automata-Markov model. *Geocarto International, 37*(4), 1183-1202.

Aysal, F. E., Yıldırım, K., & Cengiz, E. (2022). Real-Time Application of Traffic Sign Recognition Algorithm with Deep Learning. *Journal of Materials and Mechatronics: A, 3*(2), 275-289.

Barmpoutis, P., Papaioannou, P., Dimitropoulos, K. & Grammalidis, N. (2020). A review on early forest fire detection systems using optical remote sensing. *Sensors, 20*(22), 6442.

Bilgili, E., Küçük, Ö., Sağlam, B., Baysal, İ., Durmaz, B. D. & Coşkuner, K. A. (2021). Türkiye orman ekosistemlerinde yangınların ekolojik rolü. *Ekoloji ve Ekonomi Ekseninde Türkiye'de Orman ve Ormancılık,* Ankara: Sonçağ Akademi, 75-115.

Bond, W. J., Woodward, F. I. & Midgley, G. F. (2005). The global distribution of ecosystems in a world without fire. *New Phytologist, 165*(2), 525-538.

Cao, L., Shen, Z. & Xu, S. (2024). Efficient forest fire detection based on an improved YOLO model. *Visual Intelligence, 2*(1), 20.

Cengiz, E., Kelek, M. M., Oğuz, Y. & Yılmaz, C. (2022). Classification of breast cancer with deep learning from noisy images using wavelet transform. *Biomedical Engineering/Biomedizinische Technik*, *67*(2), 143-150.

Cheknane, M., Bendouma, T., & Boudouh, S. S. (2024). Advancing fire detection: two-stage deep learning with hybrid feature extraction using faster R-CNN approach. Signal, Image and Video Processing, 1-8.

Chen, Y., Zhang, Y., Xin, J., Wang, G., Mu, L., Yi, Y., ... & Liu, D. (2019). UAV image-based forest fire detection approach using convolutional neural network. In *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)* (pp. 2118-2123). IEEE.

Chuvieco, E., Aguado, I., Salas, J., García, M., Yebra, M. & Oliva, P. (2020). Satellite remote sensing contributions to wildland fire science and management. *Current Forestry Reports*, *6*, 81-96.

Emmy Prema, C., Vinsley, S. S. & Suresh, S. (2018). Efficient flame detection based on static and dynamic texture analysis in forest fire detection. *Fire Technology*, *54*, 255-288.

FAO, (2020). Global forest resources assessment 2020, global forest resources assessment 2020. FAO Rome.

Gibson, L., Lee, T. M., Koh, L. P., Brook, B. W., Gardner, T. A., Barlow, J., ... & Sodhi, N. S. (2011). Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature*, *478*(7369), 378-381.

Girshick, R., Donahue, J., Darrell, T. & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 580-587).

Glasspool, I. J., Edwards, D. & Axe, L. (2004). Charcoal in the Silurian as evidence for the earliest wildfire. *Geology, 32*(5), 381-383.

Han, X. F., Jin, J. S., Wang, M. J., Jiang, W., Gao, L. & Xiao, L. P. (2017). Video fire detection based on Gaussian Mixture Model and multi-color features. *Signal, Image and Video Processing, 11*, 1419-1425.

Jeong, J., Park, H. & Kwak, N. (2017). Enhancement of SSD by concatenating feature maps for object detection. arXiv preprint arXiv:1705.09587.

Kelek, M. M., Cengiz, E., Oğuz, Y. & Yönetken, A. (2021). Examination and classification of mammography images with the RLBP method. *AKU-IJETAS, 4*, 59-64.

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

Li, J., Xu, R. & Liu, Y. (2023). An improved forest fire and smoke detection model based on yolov5. *Forests, 14*(4), 833.

Lin, J., Lin, H. & Wang, F. (2023). A semi-supervised method for real-time forest fire detection algorithm based on adaptively spatial feature fusion. *Forests, 14*(2), 361.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y. & Berg, A. C. (2016). SSD: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing.

Lu, K., Xu, R., Li, J., Lv, Y., Lin, H. & Liu, Y. (2022). A vision-based detection and spatial localization scheme for forest fire inspection from UAV. *Forests, 13*(3), 383.

Mahmoud, M. A. I. & Ren, H. (2019). Forest fire detection and identification using image processing and SVM. *Journal of Information Processing Systems, 15*(1), 159-168.

Mukhiddinov, M., Abdusalomov, A. B. & Cho, J. (2022). A wildfire smoke detection system using unmanned aerial vehicle images based on the optimized YOLOv5. *Sensors, 22*(23), 9384.

Nesha, M. K., Herold, M., De Sy, V., Duchelle, A. E., Martius, C., Branthomme, A., ... & Pekkarinen, A. (2021). An assessment of data sources, data quality and changes in national forest monitoring capacities in the Global Forest Resources Assessment 2005–2020. *Environmental Research Letters*, *16*(5), 054029. https://doi.org/10.1088/1748-9326/abd81b

Özer, T., Akdoğan, C., Cengiz, E., Kelek, M. M., Yildirim, K., Oğuz, Y. & Akkoç, H. (2022, September). Cherry tree detection with deep learning. In *2022 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1-4). IEEE.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., ... & Hayes, D. (2011). A large and persistent carbon sink in the world’s forests. *Science*, *333*(6045), 988-993. https://doi.org/10.1126/science.1201609

Redmon, J. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Ren, S. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. arXiv preprint arXiv:1506.01497.

Sasaki, N. (2021). Timber production and carbon emission reductions through improved forest management and substitution of fossil fuels with wood biomass. *Resources, Conservation and Recycling, 173*, 105737.

Sivrikaya, F., Günlü, A., Küçük, Ö. & Ürker, O. (2024). Forest fire risk mapping with Landsat 8 OLI images: Evaluation of the potential use of vegetation indices. *Ecological Informatics, 79,* 102461.

Song, Y. & Fu, Z. (2018). Uncertain multivariable regression model. *Soft Computing, 22*, 5861-5866.

Sullivan, F. B., Ducey, M. J., Orwig, D. A., Cook, B. & Palace, M. W. (2017). Comparison of lidar-and allometry-derived canopy height models in an eastern deciduous forest. *Forest Ecology and Management*, *406*, 83-94.

Talaat, F. M. & ZainEldin, H. (2023). An improved fire detection approach based on YOLO-v8 for smart cities. *Neural Computing and Applications, 35(28)*, 20939-20954.

Tan, M., Pang, R. & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10781-10790).

Thonicke, K., Venevsky, S., Sitch, S. & Cramer, W. (2001). The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model. *Global Ecology and Biogeography, 10(6)*, 661-677.

Wang, R., Derdouri, A. & Murayama, Y. (2018). Spatiotemporal simulation of future land use/cover change scenarios in the Tokyo metropolitan area. *Sustainability, 10(6),* 2056.

Xie, Z., Song, W., Ba, R., Li, X. & Xia, L. (2018). A spatiotemporal contextual model for forest fire detection using Himawari-8 satellite data. *Remote Sensing, 10(12),* 1992.

Yacouby, R. & Axman, D. (2020, November). Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. In *Proceedings of the first workshop on evaluation and comparison of NLP systems* (pp. 79-91).

Yadav, V. S., Yadav, S. S., Gupta, S. R., Meena, R. S., Lal, R., Sheoran, N. S. & Jhariya, M. K. (2022). Carbon sequestration potential and CO2 fluxes in a tropical forest ecosystem. *Ecological Engineering, 176*, 106541.

Yang, Z., Shao, Y., Wei, Y. & Li, J. (2024). Precision-Boosted Forest Fire Target Detection via Enhanced YOLOv8 Model. *Applied Sciences, 14(6),* 2413.

Yıldız, E. & Serttaş, F. (2023). Farklı Yapay Sinir Ağları Yöntemlerini Deneysel Olarak Ölçülen Güneş Işınım Tahminine Etkisi. *Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi, 23(4)*, 962-972.

Zhou, K., Cao, L., Liu, H., Zhang, Z., Wang, G. & Cao, F. (2022). Estimation of volume resources for planted forests using an advanced LiDAR and hyperspectral remote sensing. *Resources, Conservation and Recycling*, *185*, 106485.

Zou, Z., Chen, K., Shi, Z., Guo, Y. & Ye, J. (2023). Object detection in 20 years: A survey. *Proceedings of the IEEE, 111(3)*, 257-276.

CHAPTER IV

Evaluating the Infestation Assessment of the Pine Processionary Moth using Remote Sensing Techniques

Sude TÜYSÜZOĞLU[1]
Gonca Ece ÖZCAN[2]

Introduction

Pine processionary moth (PPM) [*Thaumetopoea pityocampa* (Den. & Schiff.) / *Thaumetopoea wilkinsoni* (Tams)] is a species that causes damage, especially in pine forests (İpekdal & et al., 2015). In regions exhibiting severe forest degradation, dead stands of leafless trees may emerge, along with damage observed on individual trees (Lambers, Chapin & Pons, 1998). The species, widely dispersed over Europe, Asia, Africa, and North America, is recognized for causing damage in the coastal regions of the Mediterranean, Aegean, Marmara, and Black Sea in Türkiye, as well as in the south-facing

[1] Orman Mühendisi., Kastamonu Üniversitesi, Orman Fakültesi, Orman Mühendisliği Bölümü, Kastamonu/Türkiye, sudetuysuzoglu@ilcoud.com
[2] Doç. Dr., Kastamonu Üniversitesi, Orman Fakültesi, Orman Mühendisliği Bölümü, Kastamonu/Türkiye, Orcid: 0000-0003-0141-1031, goncaece@kastamonu.edu.tr

areas of Central Anatolia. *T. pityocampa* has distribution in the North Aegean and Thrace, whereas *T. wilkinsoni* is located in the Mediterranean, Aegean, and Black Sea regions. A potential hybrid of the two species was discovered in Türkiye (İpekdal & et al., 2015). The primary hosts include *Pinus brutia*, *P. halepensis*, *P. mugo*, *P. radiata, P. nigra, P. pinaster, and P. sylvestris* (Cayuela, Hódar & Zamora, 2011; Hódar, Castro & Zamora, 2003; Petrakis & et al., 2005; Stastny & et al., 2006), alongside *Cedrus atlantica*, *C. deodara, C. libani, and Juniperus excelsa*. It has also been noted to inflict injury on *P. heldreichii, P. taeda, P. pinea, P. strobus, P. elliottii, Pseudotsuga menziesii* (Avtzis, 1986; Cayuela, Hódar & Zamora, 2011; Kanat, Sivrikaya & Serez, 2002; Petrakis & et al., 2005). In Türkiye, it generally causes damage in pine forests and rarely in cedar forests (Can & Özçankaya, 2003).

PPM lays their eggs by aggregating two or more needles, typically commencing from the lower branches towards the top of the tree, in cylindrical clusters around the needles, like corn cobs, then surround the egg clusters with a protective covering. The color of the eggs is yellowish-white. There are 10 eggs arranged transversely in the cobs and 20-30 eggs distributed longitudinally, resulting in a total of 150-300 eggs. Egg dimensions may fluctuate. They typically measure 25-40 mm in width and around 5 mm in height. Upon completion of the 25-40 day incubation period, the caterpillars emerging from the eggs possess a disproportionately big head relative to their body and exhibit 16 legs, 6 of which are toroidal (Romanyk & Cadahía, 1992; Schmidt, Tanzen & Bellin, 1999; EPPO, 2004; Mirchev & et al., 2007). The species has five larval stages, with the initial stage exhibiting a dull green color. Fifth-stage larvae measure 35 to 45 mm in length. Each body segment is covered with hairs that possess allergenic characteristics (EPPO, 2004).

The eggs laid on pine needles hatch in about two weeks, and the emerging larvae start to feed on the needles nearby. Immediately after this feeding stage, the larvae start to form bags. As the larvae start to grow, they continue to feed on the other shoots around them. The bags that are formed continue to grow during these stages, and they are usually positioned towards the ends of the shoots. The caterpillars usually feed at night. It has been observed that they feed during the day in overcast weather. They do not leave their nests when the temperature drops below 6°C (Öymen, 1990; Çanakçıoğlu & Mol, 2000)

Adults have dull bluish, ventral, and dorsal hairs that vary from gray to black and white to yellow. The pupae of the species are oval. They are about 20 mm long and reddish brown (Romanyk & Cadahía, 1992). The wingspan of male adults is smaller than females, and both sexes have sharp-tipped protrusions on their thorax (EPPO, 2004). In adult individuals belonging to this genus, the web parts are dull, and their abdomens are covered with scales. Males are weaker than females, and the ends of their bodies are covered with hair. They have larvae with 8 pairs of legs and allergenic microscales on their dorsals (Çanakçıoğlu & Mol, 2000). The pine insect has one generation per year. It usually causes an epidemic every 6-8 years (Jacquet & et al., 2013).

Insects pose the greatest threat to the survival, production, and continuity of Turkish forests (Onaran & Katı, 2010). The PPM, renowned for its vast damage to trees in Türkiye for years, is present along the entire coastal strip (İpekdal & Çağlar, 2011). Pine needles are eaten by insect larvae (Kanat, Sivrikaya & Serez, 2002; Kerdelhué & et al., 2009), causing tree leaf loss. They can kill the tree in advanced stages. Needle leaf eating, they cause much leaf loss and limited tree development (Jacquet & et al., 2013).

The PPM induces economic detriment by reducing tree diameter and height, health issues stemming from its allergic properties, and aesthetic concerns resulting from foliar loss (Mendel, 1990). In old forests, trees rarely die, but losses can add up. It has been determined that trees with significant leaf loss have a diameter of 24%, a height of 36%, and a volume of 52% (Carus, 2004). Research indicates a 38.2% reduction rate in Calabrian pine forests (Kanat, Sivrikaya & Serez, 2002).

In controlling these pests, regional assessment of the damage is crucial for deciding the appropriate measures to implement. Nonetheless, the on-site identification of damage inflicted by the PPM on forests is a labor-intensive and expensive endeavor. Consequently, alternative procedures that are cost-effective and maintain an acceptable degree of accuracy are required. Remote sensing techniques are efficiently employed for the detection and monitoring of insect damage. The identification of insect damage by remote sensing data is sensitive to the extent of leaf damage, resulting in leaf discoloration and eventual tree mortality (Wulder & et al., 2006). Each satellite image displays varying degrees of insect damage (White, Wulder & Grills, 2006). Landsat, which is medium-resolution satellite imagery, is proficiently utilized in studies of insect damage detection owing to its extensive area coverage, multispectral bands, and 16-day temporal resolution (Collins & Woodcock, 1996; Skakun, Wulder & Franklin, 2003).

Study Area

The study area is the black pine forests in Kastamonu Regional Directorate of Forestry (RDF), İnebolu Forest Enterprise (FE). In the region with the Black Sea climate, summers are dry and hot; winters are warm and rainy. A transition to a continental climate is observed in the inland areas. The annual average temperature is 13.2 C°, and the total rainfall is 1019.77 mm. The area of İnebolu FE is 66,468.70 hectares in total, 47,180.10 hectares of which are forests and 19,288.60 hectares of which are open areas.

Database

Records regarding the infestation caused by the PPM in the black pine forests in Kastamonu RDF, İnebolu FE in 2016, when it caused intense damage, were obtained from the "Forest Pest Control Project" tables and the field survey. Using these records, 30 forest stands were determined from the Çkbc2 stand type infected by the PPM in the study area, and 37 forest stands were determined from the Çkbc2 stand type with uninfected. Figure 1 presents the bags and adult specimens of the PPM found on the black pine tree in the research area this year.

Figure 1. The PPM and image of the processionary moth on the tree

Method

Thirty compartments of the Çkbc2 stand type were identified as infected by the PPM, while thirty-seven were found to be clear of the pest. These areas were processed into the stand map in the ArcGIS environment. Landsat 8 OLI satellite imagery from 2016 was utilized to compute Normalized Vegetation Index (NDVI) values in areas infected and uninfected by PPM damage. Landsat 8 OLI satellite imagery was acquired at no cost from the website (USGS, 2016).

Table 1: Landsat 8 OLI imagery information

Acquisition Date	Cloud cover (%)	Bands	Wavelength (µm)	Resolution (m)
19/09/2016	1,04	Blue	0.45-0.51	30
		Green	0.53-0.59	
		Red	0.64-0.67	
		NIR	0.85-0.88	
		SWIR 1	1.57-1.65	
		SWIR 2	2.11-2.29	

Landsat 8 OLI satellite images were preprocessed to make them ready for analysis. First, the images were atmospherically corrected, and the digital number values of the NDVI bands were transformed into reflectance values. In the study area, 30 and 37 infected and uninfected forest stands were satellite-imaged. NDVI was determined from reflectance measurements. NDVI values were calculated for 67 forest stands using the formula below.

NDVI = (NIR – RED) / (NIR + RED) (Formula 1)

The minimum, average, maximum, and total NDVI values for each stand were calculated, considering the variability in the area of each stand and the number of pixels included within it. The compliance of the obtained NDVI values with normal distribution was tested using the one-sample Kolmogorov-Smirnov test ($p>0.05$). After the normality control, the statistical differences between the NDVI values of the stands with insect damage and those without were tested using the independent t-test for data showing normal distribution in NDVI values. We used IBM SPSS version 23 for all of our statistical studies.

Results and Discussion

NDVI was utilized to assess the damage inflicted by the PPM. NDVI images were utilized to assess the damage inflicted by the PPM on black pine trees in the Kastamonu RDF, İnebolu FE, in 2016. NDVI images of the Çkbc2 stand type with infected and uninfected are given in Figure 2.

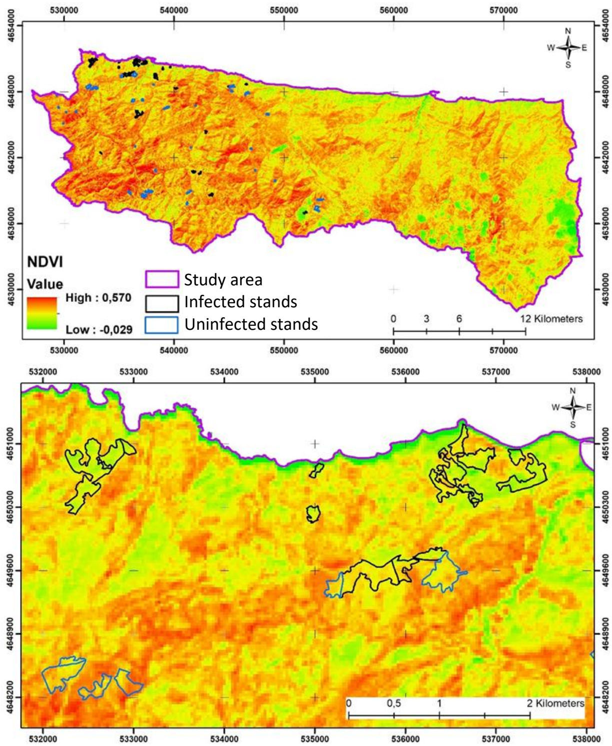

Figure 2. NDVI images of infected and uninfected forest stands (2016 Çkbc2)

The differences between the minimum, average, and maximum NDVI values obtained from a total of 67 forest stands in 2016, where infestation was observed (30 stands) and not observed (37 stands), were statistically evaluated. The normality of the data was checked using the Kolmogorov-Smirnov test ($p > 0.05$). According to the analysis, minimum NDVI, maximum NDVI, and average NDVI values show normal distribution (Table 2).

Table 2: Normality control of minimum, average, maximum, and total NDVI values according to the Kolmogorov-Smirnov test

NDVI	N	Average	Standard deviation	Minimum	Maximum	P*
Minimum NDVI	67	0.2582	0.06376	0.03	0.36	0.255
Maximum NDVI	67	0.3843	0.05212	0.21	0.50	0.992
Average NDVI	67	0.3170	0.04999	0.15	0.42	0.727

*$P < 0.05$

In this study, minimum, maximum, and average NDVI values in infected stands by PPM were calculated as 0.2085, 0.3451, and 0.2727, respectively. In contrast, minimum, maximum, and average NDVI values in infected stands were calculated as 0.2985, 0.4161, and 0.3528, respectively. Minimum, maximum, and average NDVI values in uninfected stands were higher than NDVI values in infected stands (Table 3). Accordingly, the differences between minimum NDVI, maximum NDVI, and average NDVI values showing normal distribution in stands infected and uninfected stands were determined by independent t-test. Minimum NDVI, maximum NDVI, and average NDVI values in infected and uninfected stands were statistically different (p<0.05) (Table 3).

Table 3: Independent t-test results of the differences between minimum, maximum, and average NDVI values in infected and uninfected stands

NDVI	Infestation	n	Average	Standard deviation	p
Minimum NDVI	Infected stands	30	0.2085	0.05962	0.000
	Uninfected stands	37	0.2985	0.02942	
Maximum NDVI	Infected stands	30	0.3451	0.04189	0.000
	Uninfected stands	37	0.4161	0.03533	
Average NDVI	Infected stands	30	0.2727	0.03668	0.000
	Uninfected stands	37	0.3528	0.02323	

*P<0.05

In this study, minimum, maximum, and average NDVI values were found to be statistically different in infected and uninfected stands. In a study conducted in the same 70 stands in the Calabrian pine forests of the Kahramanmaraş RDF, Elmalar FE in 2016, when there was intense PPM damage, and in 2022, when no damage was observed, statistically significant differences were found between the NDVI data (Özcan & Sivrikaya, 2022). In this study, it was observed that the minimum, maximum, and average NDVI values in infected stands with PPM were higher than in uninfected stands. Similar results were found in a study conducted with the same pest in different regions and tree species (Özcan & Sivrikaya, 2022). NDVI values are lower in years with infected stands due to the decrease in NIR band values (Junttila & et al., 2015). In a study conducted to determine leaf loss caused by Lymantria dispar in broadleaved forests, leaf losses occurring in two consecutive years were evaluated. In the study, NDVI and EVI indices were widely used. It was determined that NDVI performed significantly better in determining leaf fall in large areas (>0.6 km2) in MODIS 250 m data. In addition, they recommended MODIS due to its higher spatial resolution and the results obtained compared to alternative indices such as NDVI (De Beurs & Townsend, 2008). NDVI is more effective in differentiating moderate damage from low damage in broadleaved trees (Spruce & et al., 2011). In addition, Jepsen & et al. (2009) stated that NDVI is more reliable than MODIS in long-term monitoring due to significant cloud cover during the periods when leaf fall occurs.

The damage caused by the PPM can have considerable negative impacts on the forest environment. Assessing the harm inflicted by this species on forests using on-site field surveys is laborious, time-consuming, and expensive. Consequently, assessing the harm inflicted by the insect and monitoring it through remote sensing techniques will aid in mitigating these adverse effects. This study assessed the infected stands by the insect using remote sensing techniques. The results will help in understanding the infestation of this species across extensive regions with reduced work and expense, hence enhancing effectiveness in pest management. It is crucial for mitigating harm to forest resources and implementing appropriate measures for ecosystem sustainability.

Acknowledge

This research was funded by the TÜBİTAK 2209-A University Students Research Projects Support Program (1919B012201182). We would like to express our sincere gratitude to the Kastamonu Regional Directorate of Forestry, for providing the data used in this study.

References

Avtzis, N. (1986). Development of *Thaumetopoea pityocampa* Schiff. (Lepidoptera: Thaumetopoeidae) in relation to food consumption. *Forest Ecology and Management 15* (1), 65-68.

Can, P. & Özçankaya, M. (2003). Ege Bölgesi ağaçlandırma alanlarında çam kese böceği (*Thaumetopoea pityocampa* Schiff.) yumurta parazitoitlerinin belirlenmesi, *Ege Ormancılık Araştırma Müdürlüğü Teknik Bülten, 22*: 42

Çanakcıoğlu, H. & Mol, T. (2000). Tohum ve kültür zararlıları. İstanbul Üniversitesi Orman Fakültesi Yayınlar Rektörlük, 4210: 198-202.

Carus, S. (20049. Impact of defoliation by the pine processionary moth (*Thaumetopoea pityocampa*) on radial, height and volume growth of Calabrian pine (*Pinus brutia*) trees in Turkey. *Phytoparasitica, 32* (5), 459-469

Cayuela, L., Hódar, JA. & Zamora, R. (2011). Is insecticide spraying a viable and cost-efficient management practice to control pine processionary moth in Mediterranean woodlands? *Forest Ecology and Management, 261* (11), 1732-1737.

Collins, J.B. & Woodcock, CE. (1996). An assessment of several linear change detection techniques for mapping forest mortality using multitemporal landsat TM data. *Remote Sensing of Environment, 56*, 66–77.

De Beurs, K. & Townsend, P. (2008). Estimating the effect of gypsy moth defoliation using MODIS. *Remote Sensing and Environment, 112*, 3983-3990

EPPO, (2004). Diagnostic protocols for regulated pests protocoles de diagnostic pourles organismes reglementes. *OEPP/EPPO Bulletin, 34* (2), 271-279.

Hódar, JA., Castro, J. & Zamora, R. (2003). Pine processionary caterpillar *Thaumetopoea pityocampa* as a new threat for relict Mediterranean Scots pine forests under climatic warming. *Biological conservation 110*, (1),123-129.

İpekdal, K. & Çağlar, SS. (2011). *Thaumetopoea pityocampa* ve *T. wilkinsoni*'nin Türkiye'deki yayılışının ve melezleşmesinin moleküler yöntemlerle araştırılması. Türkiye I. Orman Entomolojisi ve Patolojisi Sempozyumu, 23-25 Kasım 2011, Antalya, Türkiye, (pp. 118-122)

İpekdal, K., Burban, C., Kerdelhue, C. & Çağlar, SS. (2015). Distribution of two pine processionary moth species in Turkey evidences a contact zone. *Turkish Journal of Zoology, 39* (5), 868-876.

Jacquet, JS., Bosc, A., O'Grady, A.P. & Jactel. H. (2013). Pine growth response to processionary moth defoliation across a 40-year chronosequence. *Forest Ecology and Management*, *293*, 29-38.

Jepsen, U., Hagen, S., Ims, R. & Yoccoz, N. (2009). Climate change and outbreaks of the geometrids *Operophtera brumata* and *Epirrita autumnata* in subarctic birch forest: evidence of a recent outbreak range expansion. *Journal of Animal Ecology, 77*, 257-264.

Junttila, S., Kaasalainen, S., Vastaranta M., Hakala, T., Nevalainen, O. & Holopainen, M. (2015). Investigating bi-temporal hyperspectral lidar measurements from declined trees– experiences from laboratory test. *Remote Sensing*, *7,* 13863–13877.

Kanat, M., Sivrikaya, F. & Serez, M. (2002). A research on damage of pine processionary moth (*Thaumetopoea pityocampa* Schiff.) on *Pinus brutia* Ten. trees, and the effect of tending activities on the diameter increment of *Calabrian pine* in Kahramanmaraş. Pine Processionary Moth Symposium, April 24–25, 2002, Kahramanmaraş, Turkey (pp 44-51)

Kerdelhué, C., Zane, L., Simonato, M., Salvato, P., Rousselet, J., Roques, A. & Battisti, A. (2009). Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology, 9* (1), 220

Lambers, H., Chapin III F.S & Pons, T.L. (1998). Plant physiological ecology (Book), Second edition, Springer, NewYork, p. 603.

Mendel, Z. (1990). On the origin of the pine processionary caterpillar, *Thaumetopoea wilkinsoni* Tams (Lep.: Thaumetopoeidae) in Israel. *Journal of Applied Entomology, 109*, (3), 311-314.

Mirchev, P., Tsankov, G., Avci, M. & Matova, M. (2007). Study of some aspects of ecology of pine processionary moth, *Thaumetopoea pityocampa* (Den. & Schiffe.) (Lep., Thaumetopoeidae) and its egg parasitoids in Turkey. *Silva Balcanica 8* (1), 66-78.

Onaran, MA. & Katı, M. (2010). Çam kese böceği (*Thaumetopoea pityocampa* Shiff) ile biyolojik mücadele. Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 12 (2), 21-27.

Öymen, T. (1990). Türkiye'de iğne yapraklı ağaçlarda zarar yapan Lepidoptera türleri. İstanbul Üniversitesi Orman Fakültesi Dergisi, 40 (3), 59-66.

Özcan, GE. & Sivrikaya, F. (2022). Determining infestation of Pine Processionary Moth using remote sensing. Intercontinental Geoinformation Days, 4, 99-102.

Petrakis, PV., Roussis, V., Papadimitriou, D., Vagias, C. & Tsitsimpikou, C. (2005). The effect of terpenoid extracts from 15 pine species on the feeding behavioural sequence of the late instars of the pine processionary caterpillar *Thaumetopoea pityocampa*. *Behavioural Processes*, *69* (3), 303-322.

Romanyk, N. & Cadahía, N. (1992). Plagas de pinos y otras coníferas. Plagas de insectos en las masas forestales espanolas. 2a ed. Madrid: Ministerio de Medio Ambiente.

Schmidt, G., Tanzen, E. & Bellin, S. (1999). Structure of egg-batches of *Thaumetopoea pityocampa* (Den. and Schiff.) (Lep., Thaumetopoeidae), egg parasitoids and rate of egg parasitism on the Iberian Peninsula. *Journal of Applied Entomology 123* (8), 449-458.

Skakun, RS., Wulder, MA. & Franklin, SE. (2003). Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sensing of Environment*, *86*, 433–443.

Spruce, J., Sader, S., Ryan, R., Smoot, J., Kuper, P. & Ross, K. (2011). Assessment of MODIS NDVI time series data products for detecting forest defoliation by gypsy moth outbreaks. *Remote Sensing and Environment*, *115*, 427-437.

Stastny, M., Battisti, A., Petrucocco-Toffolo, E., Schlyter, F. & Larsson, S. (2006). Host-plant use in the range expansion of the pine processionary moth, *Thaumetopoea pityocampa*. *Ecological entomology, 31* (5), 481-490.

Townsend, P., Singh, A., Foster, J., Rehberg, N., Kingdon, C. & Eshleman, K. (2012). A general Landsat model to predict canopy defoliation in broadleaf deciduous forests. *Remote Sensing and Environment, 119,* 255-265.

U.S. Geological Survey (2016). https://earthexplorer.usgs.gov/. Accessed 25 March 2022.

White, J.C., Wulder, M.A. & Grills, D. (2006). Detecting and mapping mountain pine beetle red-attack damage with SPOT-5 10-m multispectral imagery. *BC Journal of Ecosystems and Management*, *7*, 105–118.

Wulder, M.A., Dymond, C.C., White, J.C., Leckie, D.G. & Carroll, A.L. (2006). Surveying mountain pine beetle damage of forests: a review of remote sensing opportunities. *Forest Ecology and Management*, *221*, 27−41.

CHAPTER V

Assessment of Damage Status of *Ips sexdentatus* Utilizing Remote Sensing Data

Cihan AK[1]
Fatih SİVRİKAYA[2]

Introduction

Insects are a crucial component of the ecological balance of forests, contributing to ecological diversity. The dynamic forest ecosystem is influenced by bark beetle infestations, whether on a small or major scale (Özcan, 2017; Özcan et al., 2022). In recent years, outbreaks of bark beetles (Coleoptera: Curculionidae, Scolytinae), which result in tree mortality increased by global climate change, have dramatically grown worldwide (Hlásny et al., 2019; Hlásny et al., 2021; Sommerfeld et al., 2021). Forecasts indicate that the outbreaks will continue to increase (Evangelista et

[1] Orman Mühendisi, Kastamonu Üniversitesi, Orman Fakültesi, Orman Mühendisliği Bölümü, Kastamonu/Türkiye, akcihan823@gmail.com
[2] Prof. Dr., Kastamonu Üniversitesi, Orman Fakültesi, Orman Mühendisliği Bölümü, Kastamonu/Türkiye, Orcid: 0000-0003-0860-6747, fsivrikaya@kastamonu.edu.tr

al., 2011; Seidl et al., 2014). Sixteen species of the genus Ips, a taxonomically classified group of bark beetles, are found in the forests of Eurasia (Cognato & Felizet, 2000).

Ips sexdentatus Boerner (Coleoptera: Curculionidae: Scolytinae) is a native species found in Turkish forests, causing harm to spruce, fir, and pine trees (Bernhard, 1935; Oymen, 1992; Özcan, Eroğlu & Alkan-Akıncı, 2011; Sivrikaya et al., 2022). *I. sexdentatus* is a beetle that falls under the order Coleoptera and is classified within the subfamily Scolytinae of the family Curculionidae. *I. sexdentatus*, measuring between 5.5 and 8.0 mm, stands out as a relatively large species within the bark beetle category (Çanakçıoğlu & Mol, 1998; Güzel, 2018). *I. sexdentatus* typically produces two generations annually, with flight times varying based on altitude and climatic conditions (Yüksel & Akbulut, 2005). *I. sexdentatus* is a secondary pest species that favors trees weakened by various factors such as snow breakage, storm overturns, air pollution, and water stress for reproduction. Nonetheless, in instances of nutrient deficiency and over-reproduction, it may also harm healthy trees (Yüksel & Akbulut, 2005). *I. sexdentatus* inflicts damage on coniferous trees across a vast region of the world, notably in Europe, Russia, the Caucasus, Asia, Siberia, Korea, and Japan, as well as on spruce in Türkiye and Georgia (Güzel, 2018).

I. sexdentatus is generally considered a secondary pest that slows down growth and causes growth loss. However, it reproduces rapidly under suitable conditions, especially in newly formed young stands (Beşceli & Ekici, 1969), can become a primary pest threatening the entire forest, and can destroy the entire stand. In this respect, bark beetle has great importance in terms of economic loss in forestry activities in Türkiye. Although *I. sexdentatus* is not generally considered an aggressive species, increasing populations of more aggressive Tomicus and other Ips species can attack and kill healthy trees, especially after fires. This species favors weak, stressed, and deceased trees in endemic populations (Gil & Pajares, 1986). It can also infest healthy trees and initiate epidemics when appropriate hosts are present (Raffa & Berryman, 1983).

I. sexdentatus is among the largest species within the Ips genus and is recognized for causing harm to spruce and predominantly black pine species in Türkiye. Spruce is the tree species most adversely impacted in Türkiye. However, it is also known to inflict significant harm to *Pinus nigra*, *P. brutia*, and *P. sylvestris* species. In 2021, Calabrian pine constituted 22.7% of our total forest area, black pine 18.3%, Scots pine 6.1%, and spruce 1.6%. Consequently, 48.7% of our forests consist of tree species infected by *I. sexdentatus*. *I. sexdentatus* represents a significant biological danger to these tree species, which inhabit about fifty percent of the country's forests (URL, 1).

Türkiye's entire forest area comprises 27.6% of the nation's overall land area. Moreover, it is a species with significant afforestation potential in Türkiye's steppe and semi-arid regions. Sixty-four percent of Kastamonu province, significant in Türkiye for its forest area, forestry operations, and forestry economy, is forested. In the region's forests, *I. sexdentatus*, *Pityokteines curvidens*, *Ips acuminatus*, *Tomicus piniperda*, *Cryphalus piceae*, and *Thaumetopoea pityocampa*, which caused considerable damage primarily to coniferous species, have been identified, and it has been noted that bark beetles cause severe harm. Between 2007 and 2011, a total of 100,173.9 hectares of land were infected, including 24% in 2011, 21% in 2010, 19% in 2008, 18% in 2007, and 17% in 2009 (Bayırcık, 2018).

In addition to biotechnical control against bark beetles using pheromone traps, removing dead or dried trees near healthy stands are important measures. The locational determination of damaged trees and stands is of critical importance in terms of measures to be taken in pest control. However, in order to detect the damage caused by *I. sexdentatus* to forests on site, much time and high cost are required. Therefore, alternative methods with sufficient accuracy and lower cost are needed. Remote sensing (RS) techniques are efficiently used for the identification and monitoring of *I. sexdentatus* damage (Özcan & Sivrikaya, 2022).

The degree of tree mortality and alterations in leaf pigmentation influence the identification of insect damage by RS (Wulder et al., 2006). Satellite imagery reveals varied degrees of insect damage (White, Wulder & Grills, 2006). Medium-resolution Landsat satellite imagery effectively identifies and delineates insect damage (Özcan & Sivrikaya, 2022; Skakun, Wulder & Franklin, 2003). Insect damage can be assessed through the desiccation of tree leaves with RS data (Entcheva, Cibula & Carter, 1996), hence allowing for the use of tree reflectance values and NDVI. Vegetative indices, particularly NDVI, have proven effective in ecological applications as they directly quantify vegetative productivity (Walter & Platt, 2013). NDVI utilizes the principle that leaf chlorophyll absorbs red light and correlates with chlorophyll concentration, while leaf mesophyll structure reflects near-infrared (NIR) wavelengths (Tucker, 1979; Pettorelli et al., 2005).

Satellite data is an effective tool for mapping and monitoring insect outbreaks. While there are many studies on this subject in the world, there are very few studies in Türkiye. The main purpose of this study is to determine *I. sexdentatus* damage with RS data and to test its applicability.

Study Area

Mergüze Forest Planning Unit, which is located in Kastamonu Regional Directorate of Forestry, İhsangazi Forest Enterprise, was selected as the case study area. The study area borders are at 41°13'46" N, 33°27'01" E. The area of the Mergüze Forest Planning Unit is 11164.9 ha, and approximately 62% (6887.8 ha) of the area is covered with forest areas. The productive forest area is 4616.4 ha. The main tree species in the forest ecosystem are black pine, Scots pine, oak, fir, and hornbeam. Although İhsangazi District is located in the Black Sea climate zone, it has a more difficult climate structure in terms of temperature and precipitation characteristics. The highest temperature in the history of the district was determined as 37.7 C° and the lowest temperature as -26.9 C°. While an increase in the amount of precipitation is observed in May and June due to the effect of atmospheric conditions, this situation exhibits a decreasing trend in July and August.

Database and Method

Data from the "Forest Pest Control Project," along with field surveys conducted by the Kastamonu Regional Directorate of

Forestry (RDF), İhsangazi Forest Enterprise, Mergüze Forest Planning Unit, indicate that *I. sexdentatus* damage its most severe harm in 2023. Consequently, based on the relevant information and field survey for 2023, a total of 60 forest stands were identified, comprising 30 from the Çkbc2 stand type, both with and without *I. sexdentatus* damage. The stand types were analyzed using the digital stand map acquired from the Kastamonu RDF within the ArcGIS environment.

Landsat 8 OLI satellite imagery from 2023 was utilized to compute and compare NDVI values in stands with and without *I. sexdentatus* damage. The 2023 satellite imagery was obtained at no cost from the website, and the image characteristics are presented in Table 1 (USGS, 2023). Essential preliminary procedures were conducted for the analysis of the Landsat 8 OLI satellite imagery utilized in the study. The satellite image performed atmospheric correction, and the reflectance values required to calculate NDVI were determined.

In the subsequent phase, the forest stands with (30) and without (30) *I. sexdentatus* damage in the research region was incorporated into the satellite image. NDVI values were computed based on the reflectance measurements for the 60 specified forest stands. NDVI is employed to assess vegetation existence by enhancing the contrast between the near-infrared (NIR) and red bands, subsequently consolidating the data from these two bands into a singular band (Duran, 2007). The above formula was employed to compute the NDVI value.

Table 1: Fundamental characteristics of Landsat 8 OLI satellite imagery

Acquisition Date	Cloud cover (%)	Bands	Wavelength (µm)	Resolution (m)
2023/09/23	0.62	Blue	0.45-0.51	30
		Green	0.53-0.59	
		Red	0.64-0.67	
		NIR	0.85-0.88	
		SWIR 1	1.57-1.65	
		SWIR 2	2.11-2.29	

NDVI = (NIR – RED) / (NIR + RED)

The minimum, mean, and maximum NDVI values for each forest stand were computed. The normal distribution of the estimated NDVI values was assessed using the Kolmogorov-Smirnov normality test ($p>0.05$). Following the normality assessment, statistical differences in NDVI values between stands with and without *I. sexdentatus* damage and those devoid of such damage were analyzed using an independent t-test for data demonstrating normal distribution in NDVI values. Statistical analyses were performed with IBM SPSS version 23 software.

Results and Discussion

NDVI values were employed to assess *I. sexdentatus* damage utilizing Landsat 8 OLI satellite images. To illustrate the impact of *I. sexdentatus* on the black pine forest within the Mergüze Forest

Planning Unit, NDVI from the Çkbc2 forest stands, with and without bark beetle damage in 2023, are presented in Figure 1 and Figure 2.

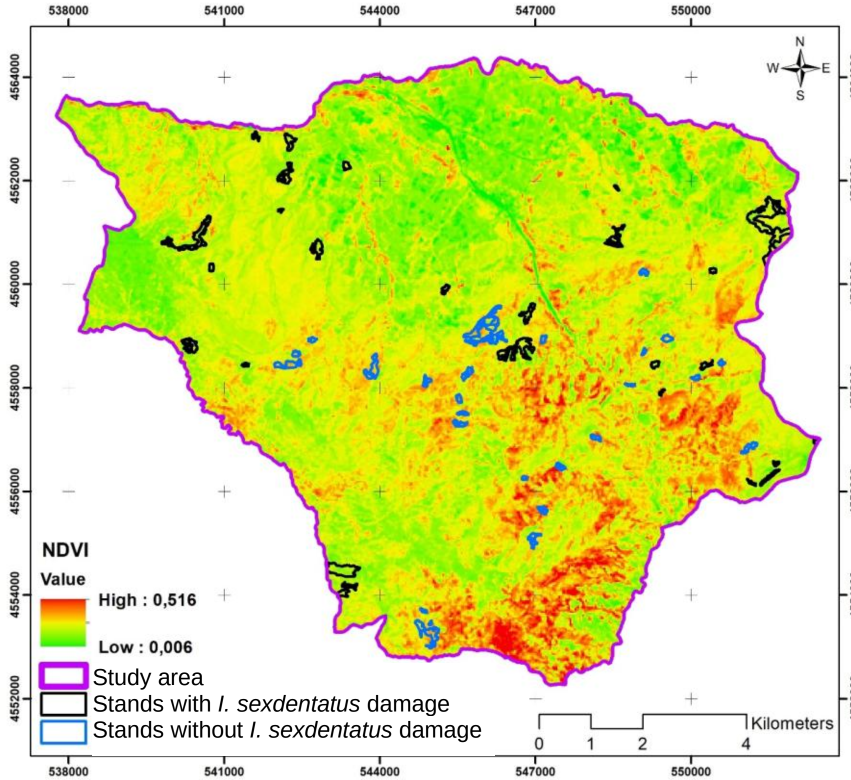

Figure 1. NDVI of Çkbc2 stands with and without I. sexdentatus damage

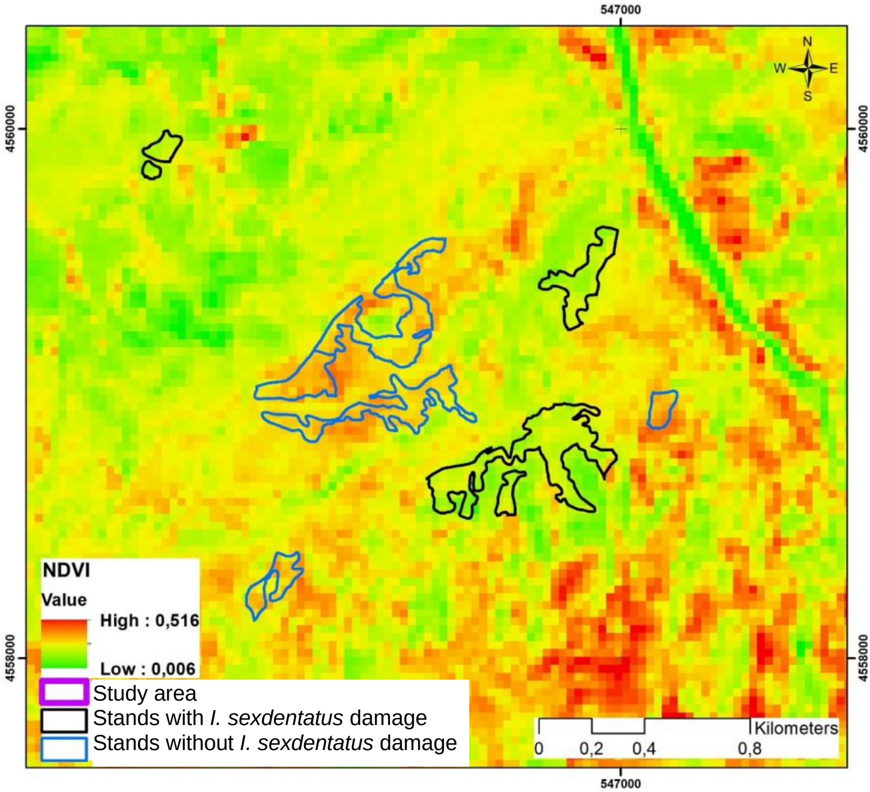

Figure 2. Detailed view of NDVI in Çkbc2 stands with and without I. sexdentatus damage

In 2023, NDVI values of stands of with and without *I. sexdentatus* damage was statistically analyzed. This investigation statistically analyzed the differences among minimum, maximum, and mean NDVI values. The research revealed that the maximum, minimum, and mean NDVI values exhibited a normal distribution ($p > 0.05$) (Table 2).

Table 2: Assessment of normality for minimum, maximum, and mean NDVI values via the Kolmogorov-Smirnov test

NDVI	N	Mean	Standard deviation	Min.	Max.	P*
Min. NDVI	60	0.1862	0.05364	0.08	0.28	0.349
Max. NDVI	60	0.2722	0.05565	0.19	0.38	0.368
Mean NDVI	60	0.2296	0.04858	0.16	0.31	0.217

*P < 0.05

In stands damaged by *I. sexdentatus*, the minimum, mean, and maximum NDVI values were calculated as 0.1386, 0.2227, and 0.1826, respectively. Conversely, in stands without damage by *I. sexdentatus*, the minimum, mean, and maximum NDVI values were recorded as 0.2338, 0.3216, and 0.2766, respectively. The minimum, mean, and maximum NDVI values in stands without damage were substantially greater than those in areas affected by bark beetle damage (Table 3). The differences among the minimum, mean, and maximum NDVI values exhibit a normal distribution in stands with and without *I. sexdentatus* damage was examined using an independent t-test. Statistical analysis indicates that the minimum, maximum, and mean NDVI values exhibit significant differences between stands with and without *I. sexdentatus* damage ($p<0.05$) (Table 3).

Table 3: Results of the independent t-test comparing the minimum, mean, and maximum NDVI values in stands affected by I. sexdentatus damage to those without such damage.

NDVI	Damage status	n	Mean	Standard deviation	P*
Min. NDVI	Stands with damage	30	0.1386	0.02510	0.000
	Stands without damage	30	0.2338	0.02322	
Max. NDVI	Stands with damage	30	0.2227	0.02425	0.000
	Stands without damage	30	0.3216	0.02569	
Mean NDVI	Stands with damage	30	0.1826	0.00966	0.000
	Stands without damage	30	0.2766	0.01188	

*P<0.05

NDVI facilitates the examination of trees' vitality by quantifying the light absorbed and reflected by trees, as well as the presence of chlorophyll. The NDVI value of stands subjected to stress, such as insect damage or desiccation, is inferior to that of healthy forests (Stoyanova et al., 2018). NDVI values are typically recognized to range from 0.55 to 0.70 for healthy trees when evaluating stand health (Yu et al., 2020; Bryk, Kołodziej & Pliszka 2021). In stands affected by bark beetle damage, the NDVI value typically falls below 0.5 (Georgiev et al., 2022a). A study in *Pinus nigra* and *P. sylvestris* stands in Bulgaria revealed that the mean NDVI value for healthy trees was 0.617, while that for dead trees was 0.343 (Georgiev et al., 2022b).

De Beurs & Townsend (2008) utilized MODIS images to delineate defoliation caused by *Lymantria dispar* in broadleaf and oak-dominated forests during two successive years. The study employed NDVI and EVI indices with the BioSIM model. NDVI results greatly surpassed EVI in detecting insect-induced defoliation using daily MODIS 250 m data. Spruce et al. (2011) indicated that NDVI is more effective in differentiating moderate damage from moderate damage in broadleaved forests. NDVI exhibits a robust positive linear connection with vegetation cover ranging from 25% to 80% (Zhang et al., 2010).

NDVI is frequently utilized to evaluate disturbances in spruce stands. According to Lausch et al. (2013), in Germany's Bavarian Forest, the Vis and NIR wavelengths are the most significant spectral indicators of spruce health. Heurich et al. (2010) employed object-oriented image analysis of spruce bark beetle infestation in Bavarian Forest to demonstrate that the NDVI can differentiate deadwood from healthy vegetation. Mišurec et al. (2016) found that NDVI showed significant forest degradation and minor physiological changes in spruce in the Ore Mountains (Czechia), even though there was only a small amount of leaf loss. In the eastern San Juan Mountains (USA), Hart & Veblen (2015) employed NDVI to distinguish between gray and green stands in beetle-infested spruce-fir.

Huo et al. (2021) utilized Sentinel-1 and Sentinel-2 imagery to detect damage to *Ips typographus* (L.) in southern Sweden in spruce forests. They used the Normalized Distance Red & SWIR (NDRS) index. NDRS recognized stressed forests with accuracy ranging from 0.80 to 0.88 prior to the attacks, 0.80 to 0.82 during early-stage infestations, and 0.81 to 0.91 in middle-stage and late-stage infestations. Jamali et al. (2023) used Sentinel-2 data from 2015 to 2021 to investigate the early identification of *Ips typographus* infestations in Sweden. Bark beetle infestations affected wavelength bands and vegetative indices (VIs). The red band (81%), SWIR1 (74%), and SWIR2 (71%) performed well, followed by the red-edge (66%), green/blue bands (63%), and NIR (33%). NDVI exhibited lower sensitivity to attack impacts compared to NDRS and NDWI.

Anees et al. (2013) detected beetle infestations in North American pine forests using MODIS. The beetles induce gradual alterations that deteriorate the NDVI time series. NDVI time series methodologies are more effective at detection than window-based techniques, as they evaluate individual pixels sequentially. Bryk, Kołodziej & Pliszka (2021) conducted a study in the spruce forest in the Bialowieza. Quantitative NDVI distributions exhibited fluctuations, with declines in the minimum, mean, and median values, as well as alterations in the forms of index value distributions. Analysis of spatial NDVI distributions indicated that a threshold NDVI value of 0.6 differentiates betGonzaween healthy and unhealthy spruce stand locations. This study demonstrated that we can effectively check the health of spruce stands using NDVI from Landsat archives with easily accessible, medium-resolution images and that spatial NDVI distributions simplify forest monitoring on a larger scale.

Gomez et al. (2020) used MODIS and Sentinel-2 data to detect bark beetles in Florida. MODIS and Sentinel-2 are capable of identifying bark beetle damage, while MODIS NDVI change detection exhibits a 30% false negative rate. Sentinel-2 NDVI products find bark beetle disturbances more accurately than MODIS change maps because they always find negative NDVI changes. Yang (2019) said that NDVI is a good way to find bark beetle gray attacks when combined with raw bands and other greenness-related VIs from Sentinel-2 with a 10-meter resolution.

Conclusion

This study evaluated the efficacy of NDVI values derived from Landsat 8 OLI satellite data in assessing damage caused by *Ips sexdentatus*. A total of 60 forest stands exhibiting with and without *I. sexdentatus* damage in the Mergüze Forest Planning Unit. The minimum, maximum, and mean NDV values of these stands were subjected to statistical comparison. The Independent t-test results indicated significant differences among the minimum, mean, and maximum NDVI values in stands with and without *I. sexdentatus* damage. These results will help with the assessment of *I. sexdentatus* damage across extensive regions more efficiently and with reduced effort, enabling the implementation of necessary actions. Early identification of bark beetle damage will greatly influence the efficacy of controlling such damage. This study will enhance the conservation and sustainability of forests.

Acknowledge

This research was funded by the TÜBİTAK 2209-A University Students Research Projects Support Program (1919B012109033). We would like to express our sincere gratitude to the Kastamonu Regional Directorate of Forestry, for providing the data used in this study.

References

Anees, A., Olivier, J. C., O'Rielly, M., & Aryal, J. (2013). Detecting beetle infestations in pine forests using MODIS NDVI time-series data. *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS* (pp. 3329-3332). 21-26 July, Melbourne, VIC, Australia.

Bayırcık, S. (2018). Kastamonu ormanlarında 2017-2018 yıllari arasında kabuk böcek zararları, *Öğrenci Bitirme Çalışması, Kastamonu Üniversitesi*.

Bernhard, R.J. (1935). Grundlagen, geschichte und aufgaben der forstwirtschaft in der Türkei. *Ankara Yüksek Ziraat Enstitüsü*, p. 377.

Beşceli O. & Ekici, M. (1969). Doğu Ladini (*Picea orientalis* L.) Mıntıkasında *Ips sexdentatus'* un Biyolojisi ve Mücadelesi, *Ormancılık Araştırma Enstitüsü Yayınları Teknik Bülten Serisi* No:32.

Bryk, M., Kołodziej, B. & Pliszka, R. (2021). Changes of Norway spruce health in the Białowieża forest (CE Europe) in 2013–2019 during a bark beetle infestation, studied with Landsat imagery. *Forests*, *12*(1), 34.

Çanakçıoğlu, H. & Mol, T. (1998). Orman entomolojisi: Zararlı ve yararlı böcekler. İstanbul Üniversitesi, *İstanbul Üniversitesi Yayınları*.

Cognato, A.I. & Felix, A. (2000). Phylogeny of Ips DeGeer species (Coleoptera: Scolytidae) inferred from Mitochondrial Cytochrome Oxidase I DNA sequence. *Molecular Phylogenetics and Evolution, 14*, 445–460.

De Beurs, K. & Townsend, P. (2008). Estimating the effect of gypsy moth defoliation using MODIS. *Remote Sens Environ 112*: 3983-3990.

Duran, C. (2007). Uzaktan algılama teknikleri ile bitki örtüsü analizi, *Doğu Akdeniz Ormancılık Araştırma Müdürlüğü DOA Dergisi*, *Sayı* 3, 45-67.

Entcheva P. K., Cibula W.G. & Carter G.A. (1996). Spectral reflectance characteristics and remote sensing detection of southern pine beetle infestations. *Eco-Informa Conference*, Nov 4–7, 1996, Lake Buena Vista, USA.

Evangelista, P. H., Kumar, S., Stohlgren, T. J. & Young, N. E. (2011). Assessing forest vulnerability and the potential distribution of pine beetles under current and future climate scenarios in the Interior West of the US. *Forest Ecology and Management*, *262*(3), 307-316.

Georgiev, G., Georgieva, M., Dimitrov, S., Iliev, M., Trenkin, V., Mirchev, P. & Belilov, S. (2022a). Remote sensing assessment of the expansion of *Ips typographus* attacks in the Chuprene Reserve, Western Balkan Range. *Forests*, *13*(1), 39.

Georgieva, M., Belilov, S., Dimitrov, S., Iliev, M., Trenkin, V., Mirchev, P. & Georgiev, G. (2022b). Application of Remote Sensing Data for Assessment of Bark Beetle Attacks in Pine Plantations in Kirkovo Region, the Eastern Rhodopes. *Forests, 13,* 620.

Gill L.,A. & Pajares, J. A. (1986*). Los Escolítidos de las Coníferas de la Península Ibérica. Instituto Nacional de*

Investigaciones Agrarias, Ministerio de Agricultura, Pescay Alimentación. ISBN 84-7498-250-2, Madrid, Spain.

Gomez, D. F., Ritger, H. M., Pearce, C., Eickwort, J. & Hulcr, J. (2020). Ability of Remote Sensing Systems to Detect Bark Beetle Spots in the Southeastern US. *Forests , 11,* 1167.

Güzel, H. (2018). *Ips sexdentatus* (Boern.)un yoğunluğu ve morfolojisi üzerine karaçam ve sarıçam meşcere özelliklerinin etkileri, Kastamonu Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Kastamonu.

Hart, S.J. & Veblen, T.T. (2015). Detection of spruce beetle-induced tree mortality using high- and medium-resolution remotely sensed imagery. *Remote Sens. Environ. 168*, 134–145.

Heurich, M., Ochs, T., Andresen, T. & Schneider, T. (2010). Object-Orientated image analysis for the semi-automatic detection of dead trees following a spruce bark beetle (*Ips typographus*) outbreak. *Eur. J. For. Res. 129*, 313–324.

Hlásny, T., K¨onig, L., Krokene, P., Lindner, M., Montagn´e-Huck, C., Müller, J. & Seidl, R. (2021). Bark beetle outbreaks in Europe: state of knowledge and ways forward for management. *Curr. For. Rep. 7* (3), 138–165.

Hlásny, T., Krokene, P., Liebhold, A., Montagne-Huck, C., Müller, J., Qin, H., Raffa, K., Schelhaas, M.J., Seidl, R., Svoboda, M. & Viiri, H. (2019). *Living with bark beetles: impacts, outlook and management options (No. 8).* European Forest Institute.

Huo, L., Persson, H. J., & Lindberg, E. (2021). Early detection of forest stress from European spruce bark beetle attack, and a new vegetation index: Normalized distance red & SWIR (NDRS). *Remote Sensing of Environment, 255,* 112240.

Jamali, S., Olsson, P. O., Ghorbanian, A. & Müller, M. (2023). Examining the potential for early detection of spruce bark beetle attacks using multi-temporal Sentinel-2 and harvester data. *ISPRS Journal of Photogrammetry and Remote Sensing, 205,* 352-366.

Lausch, A., Heurich, M., Gordalla, D., Dobner, H.-J., Gwillym-Margianto, S. & Salbach, C. (2013). Forecasting potential bark beetleoutbreaks based on spruce forest vitality using hyperspectral remote-sensing techniques at different scales. *Forest Ecol. Manag. 308*, 76–89.

Mišurec, J., Kopaˇcková, V., Lhotáková, Z., Campbell, P. & Albrechtová, J. (2016). Detection of spatio-temporal changes of Norway spruce forest stands in Ore Mountains using Landsat time series and airborne hyperspectral imagery. *Remote Sens. 8*, 92.

Oymen, T. (1992). The forest scolytidae of Turkey. *Journal of the Faculty of Forestry Istanbul University, 42,* I, 77–91.

Özcan G. E. (2017). Assessment of *Ips sexdentatus* population considering the capture in pheromone traps and their damages under non-epidemic conditions. *Sumarski list, 141* (1-2): 47-56.

Özcan, G. E. & Sivrikaya, F. (2022). Determining infestation of Pine Processionary Moth using remote sensing. *Intercontinental Geoinformation Days*, *4,* 99-102.

Özcan, G.E., Eroğlu, M. & Alkan-Akıncı, H. (2011). Use of pheromone-baited traps for monitoring *Ips sexdentatus* (Boerner)

(Coleoptera: Curculionidae) in oriental spruce stands. *Afr. J. Biotechnol. 10* (72), 16351–16360.

Özcan, G.E., Sivrikaya, F., Sakıcı, O.E. & Enez, K. (2022). Determination of some factors leading to the infestation of *Ips sexdentatus* in Crimean pine stands. *For. Ecol. Manag. 519*, 120316.

Pettorelli N., Vik J.O., Mysterud A., Gaillard J.M., Tucker C.J. & Stenseth N.C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology and Evolution, 20*: 503–510.

Raffa, K.F., Aukema, B.H., Bentz, B.J., Carroll, A.L., Hicke, J.A., Turner, M.G. & Romme, W.H. (2008). Cross-scale drivers of natural disturbances prone to anthropogenic amplification: Dynamics of biome-wide bark beetle eruptions. *Bioscience 58*, 501–517.

Seidl, R., Schelhaas, M. J., Rammer, W., & Verkerk, P. J. (2014). Increasing forest disturbances in Europe and their impact on carbon storage. *Nature Climate Change*, *4*(9), 806-810.

Sivrikaya, F., Özcan, G. E., Enez, K. & Sakici, O. E. (2022). Comparative study of the analytical hierarchy process, frequency ratio, and logistic regression models for predicting the susceptibility to *Ips sexdentatus* in Crimean pine forests. *Ecological Informatics*, *71*, 101811.

Skakun, R.S., Wulder, M.A. & Franklin S.E. (2003). Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sensing of Environment*, *86*: 433-443.

Spruce, J., Sader, S., Ryan, R., Smoot, J., Kuper, P., Ross, K., Prados, D., Russell, J., Gasser, G., McKellip, R. & Hargrove, W. (2011). Assessment of MODIS NDVI time series data products for detecting forest defoliation by gypsy moth outbreaks. *Remote Sens Environ 115*: 427-437.

Stoyanova, M., Kandilarov, A., Koutev, V., Nitcheva, O. & Dobreva, P. (2018). Potential of multispectral imaging technology for assessment coniferous forests bitten by a bark beetle in Central Bulgaria. *In MATEC Web of Conferences* (Vol. 145, p. 01005). EDP Sciences.

Tucker, C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*: 127–150.

URL 1, https://cevreselgostergeler.csb.gov.tr/orman-alanlarinin-agac-turlerine-gore-dagilimi-i-85785#_edn1 (erişim, 12.11.2023).

USGS, (2023) U.S. Geological Survey. https://earthexplorer.usgs.gov/, 2023. (accessed 3 March 2023).

Walter J.A. & Platt R.V. (2013). Multi-temporal analysis reveals that predictors of mountain pine beetle infestation change during outbreak cycles. *Forest Ecology and Management, 302*: 308–311.

White, J.C., Wulder, M.A. & Grills D. (2006). Detecting and mapping mountain pine beetle red-attack damage with SPOT-5 10-m multispectral imagery. *BC Journal of Ecosystems and Management*, *7*: 105–118.

Wulder, M., White, J., Bentz, B. & Ebata, T. (2006). Augmenting the existing survey hierarchy for mountain pine beetle red attack damage with satellite remotely sensed data. *The Forestry Chronicle 82*: 187-202.

Yang, S. (2019). Detecting bark beetle damage with Sentinel-2 multi-temporal data in Sweden. *Student thesis series INES.*

Yu, L., Zhan, Z., Ren, L., Zong, S., Luo, Y. & Huang, H. (2020). Evaluating the Potential of WorldView-3 Data to Classify Different Shoot Damage Ratios of Pinus yunnanensis. *Forests, 11*, 417.

Yüksel, B. & Akbulut, S. (2005). Doğu ladini ormanlarında *Ips sexdentatus* (Boem.)'un doğal düşmanlarının belirlenmesi. *Journal of the Faculty of Forestry, Istanbul University| İstanbul Üniversitesi Orman Fakültesi Dergisi, 55*(2), 59-70.

Zhang, T., Zhang, X., Liu, H. & Pei, X. (2010). Application of remote sensing technology in monitoring forest diseases and pests. *Plant Diseases and Pests 1*(3): 57-62.

CHAPTER VI

Time Series Analysis in the Wood Products Industry: Methodology and Application

Hakan AYDOĞAN[1]

Introduction

Wood industry holds significant importance due to its contributions to the global economy and its critical role in natural resource management. The sustainable use of forest resources is increasingly prioritized in the production of fundamental products such as timber and paper, as well as in innovative applications like bioenergy and biomaterials (He et. al., 2023; Szulecka, 2019). The rising global demand for forest products necessitates the implementation of technologies that enhance efficiency in production processes, as well as the development of strategic plans aimed at ensuring the long-term preservation of these resources. In

[1] Dr. Öğr. Üyesi/Assistant Professor, Kastamonu University, Faculty of Forestry, Department of Foreset Industry Engineering, Kastamonu/TÜRKİYE, Orcid: 0000-0001-9482-9888, hakanaydogan@kastamonu.edu.tr

this context, time series analysis offers an indispensable tool for decision-makers and researchers in the industry to understand current trends, forecast future developments based on historical data, and steer their actions toward sustainability objectives.

Time series analysis provides in-depth insights into various issues within the forest industry (Huang et. al., 2024; Kożuch et. al, 2023; Huda et. al., 2023), including monitoring supply-demand balance, analyzing price fluctuations, and understanding the sectoral impacts. Numerous factors such as seasonal variations in wood raw materials and by-products, the impact of forest fires, and environmental influences on forest assets contribute to both short- and long-term fluctuations in the forest products market. Effectively analyzing these changes necessitates an approach that not only relies on historical data but also offers robust projections for the future, which are critical for the sustainable management of forest resources. Time series models facilitate the assessment of such complex dynamics, supporting strategic decisions aimed at minimizing sectoral risks and fostering the development of innovative approaches by bridging market and environmental data.

Time series-based analyses of the forest industry are also crucial for identifying global changes and regional differences within the sector. For example, the supply-demand balance, growth dynamics, and climate change sensitivity of forest products vary widely across developed and developing economies. Given these factors, the data-driven approaches provided by time series analysis play a key role in integrating the forest industry into global sustainability efforts.

This section will comprehensively examine the importance and applications of time series analysis within the forest industry, providing a detailed exploration of methods for analyzing sector-specific variables, modeling approaches, and forecasting techniques. Aimed at developing solutions for the forest industry's variable nature, the foundations for holistic and sustainable approaches will be laid, contributing to both industry professionals and academic research.

The role and importance of time series analysis in wood products industry

Wood products industry, reliant on forests as a renewable resource, demands careful and strategic management to ensure sustainable use. In this regard, time series analysis stands out as an essential tool for understanding the industry’s dynamic structure and forecasting critical factors such as demand fluctuations, price movements, environmental influences, and climate change. Since forestry activities are directly linked to natural conditions, seasonal effects, climatic variability, and natural events play a significant role in determining the supply-demand balance of forest resources. Time series analysis is one of the foremost methods used to understand these natural and market-driven changes and to develop effective management strategies.

The importance of time series analysis in the wood industry is reflected in its role in providing forecasts aimed at sustainable resource management. The variations in the production, export, and pricing of forest products across years, seasons, and even months significantly influence decision-making processes within the sector.

For instance, periodic fluctuations in the supply of wood raw materials lead to seasonal price changes, while extreme weather events associated with climate change affect the quality and quantity of timber. By deciphering long-term trends in these factors, time series analysis helps improve understanding of the changes occurring within the forest industry. Furthermore, analyzing both long-term trends and short-term fluctuations enables industry professionals to make more informed and effective decisions.

In addition, time series analysis adds substantial value to resource management in the forest industry. Anticipating the impacts of global environmental issues such as climate crisis on forest resources allows for the development of conservation strategies and efficient utilization of natural resources. This approach is crucial not only for enhancing environmental sustainability in resource management but also for maximizing economic returns. In forest product production processes, forecasts based on seasonal data analysis prepare the ground for predicting potential risks and ensuring the efficient use of resources toward a sustainable production cycle. By revealing trends based on historical data, time series models enable the industry to address future challenges and seize emerging opportunities.

Time series analysis plays a vital role in understanding and managing the inherently dynamic nature of the forest industry. These analyses serve as foundational tools in processes such as production planning, demand forecasting, price prediction, and understanding the impacts of environmental factors. The data-driven insights provided by time series analysis contribute to a more deliberate and long-term approach to the sustainable management of forest resources, laying an essential scientific foundation for the preservation of forest ecosystems for future generations.

Time Series Analysis Methods

Introduction to time series analysis and fundamental concepts

Time series analysis involves the examination of values recorded for a specific variable over time, focusing on patterns, trends, and seasonal variations within the data. This method, widely applied across disciplines from economics to engineering, biology to environmental sciences, aims to generate insights into future outcomes based on historical data. In a dynamic sector such as the forest industry, the significance of time series analysis becomes even more evident, given the industry's susceptibility to natural cycles and economic fluctuations. Understanding the effects of various factors such as fluctuations in the supply and demand of forest products, seasonal production volumes, price movements, and climatic influences relies on the foresight that time series analysis provides.

The fundamental concepts of time series analysis are essential for understanding data characteristics and selecting appropriate analytical methods. One of the primary concepts is the presence of a **trend** in a time series. A trend represents the general direction of a series and indicates whether data values tend to increase or decrease over the long term. For example, rising household income or an increase in forest resources may correspond to an upward trend in demand for specific forest products, necessitating long-term resource planning.

In addition to trend, **seasonal fluctuations** form a crucial part of a time series; this refers to recurring variations within the series that repeat at specific intervals. In the forest industry, seasonality is particularly evident in timber production and biomass processing. Understanding seasonal fluctuations enables more efficient planning of production processes by anticipating recurring patterns.

The third fundamental component of time series analysis, **cyclical fluctuations**, refers to variations that occur over longer periods due to external factors, such as economic cycles. These fluctuations often include unpredictable changes and are influenced by factors outside the system being studied. For instance, prices of forest products may be affected by global economic cycles, underscoring the need for both short-term forecasts and long-term planning insights within the industry.

In addition to cyclical patterns, the concept of **stationarity** is critical in time series analysis. A time series is considered stationary if its statistical properties, such as mean and variance, remain constant over time. Non-stationary series, where these properties change, complicate the analysis process and may lead to misleading results if modeling is conducted without first removing trends and seasonality.

The fourth component of time series analysis involves random and irregular fluctuations. These represent unexpected variations often triggered by unforeseen events, such as natural disasters (e.g., earthquakes or fires), which can cause sudden economic shifts. Understanding these random components is essential, particularly when developing robust models capable of accounting for unexpected disruptions in the data.

The methods used in time series analysis vary depending on the structure of the series and the purpose of the analysis. One of the most widely used methods is the **ARIMA (Autoregressive Integrated Moving Average)** model, which is particularly effective for both short-term forecasting and long-term trend analysis. ARIMA models use past values of the data to predict future values, making them especially beneficial in areas such as supply-demand forecasting for forest products. For analyzing seasonal data, models with seasonal components, such as **SARIMA (Seasonal Autoregressive Integrated Moving Average)**, are preferred (Hyndman and Athanasopoulos, 2018). These models facilitate a detailed analysis of seasonal production cycles in the forest industry, supporting the development of strategies for efficient resource utilization by capturing the seasonality inherent in production processes.

These fundamental concepts and methods in time series analysis are essential for making effective and sustainable decisions in sectors like the forest industry, which are closely tied to environmental variables. Considering features such as trend, seasonality, cyclicity, and irregular fluctuations provides industry professionals and researchers with reliable, data-driven insights. This approach contributes to establishing a sustainable management model, ensuring the preservation of forest resources for future generations.

Components of time series

A time series $\boldsymbol{Y_t}$, composed of observations $\boldsymbol{x_t}$ recorded at specific times $\boldsymbol{t}$, can include several fundamental components: trend, seasonal variations, cyclical fluctuations, and irregular variations. Representing these components formally:

T_t , the trend component, which shows the long-term direction of the series (e.g., overall increase or decrease over time), S_t, the seasonal component, reflecting regular patterns that repeat over specific intervals (e.g., quarterly or annual fluctuations), C_t, the cyclical component, indicating fluctuations influenced by economic or external cycles over longer periods, which are often irregular, I_t, the irregular component, capturing random and unpredictable variations (e.g., sudden events like natural disasters).

A time series of Y_t can be expressed as:

$$Y_t = f(T_t, S_t, C_t, I_t) \tag{1}$$

Thus, in this context, a time series is composed as a function of trend, seasonality, cyclical, and irregular fluctuations. Two primary modeling approaches can be employed here: additive and multiplicative. The additive model is represented as in Equation (2), while the multiplicative model is formulated as in Equation (3).

$$Y_t = T_t + S_t + C_t + I_t \tag{2}$$

$$Y_t = T_t * S_t * C_t * I_t \tag{3}$$

Trend analysis

The trend in a time series represents a long-term directional pattern. Due to the extended nature of these periods, the trend is typically calculated from annual data. In time series applications, periods generally range between 10 to 20 time intervals, during which the data often exhibit a linear or near-linear trend.

Several methods are utilized for estimating the trend line, including the semi-average method, the simple graphical method, and the Ordinary Least Squares (OLS) method. However, in practice, the OLS method is among the most commonly employed for trend estimation. Briefly, the OLS estimation method identifies a line that best represents the distribution of data points by minimizing the sum of squared errors, thereby producing the line that best fits the data.

In a time series, let Y_t represent the observed values and $\widehat{Y}_t$ represent the corresponding values on the estimated line. The sum of squared errors (SSE) is then given by:

$$SSE = \sum_{i=1}^{n} \left(Y_t - \hat{Y}_t\right)^2 \tag{4}$$

In this case, if the independent variable X_t denotes time, the estimated trend line can be expressed as:

$$\widehat{Y}_t = \beta_{0t} + \beta_{1t} X \tag{5}$$

Here;
β_{0t} represents the intercept, or the value of the line when $X = 0$

β_{1t} represents the slope of the trend line, indicating the rate of change over time.

Example 1: Turan Wood Entegre Inc. is a company operating in the forest products sector, specializing in the production of MDF, particle board, and OSB. Table 1 below presents the MDF board production quantities for the company from 2010 to 2023. Based on this data, determine the time-dependent trend in the company's production quantities.

Table 1: MDF production quantities by year

Year	X	Production Quantities (*1000 m^3)
2010	0	500.34
2011	1	525.44
2012	2	505.37
2013	3	544.44
2014	4	566.91
2015	5	575.88
2016	6	565.11
2017	7	580.25
2018	8	584.84
2019	9	588.88
2020	10	591.34
2021	11	577.35
2022	12	595.34
2023	13	600.44

Solution:

We can use the OLS method to estimate the trend line. Accordingly, with:

$$\overline{X_t} = \frac{\sum_{t=1}^{n} X_t}{n} = \frac{91}{14} = 6.5$$

and

$$\overline{Y_t} = \frac{\sum_{t=1}^{n} Y_t}{n} = \frac{7901.93}{14} = 564.42,$$

the trend line can be estimated as:

$$\widehat{Y_t} = \beta_{0t} + \beta_{1t} X_t = 518.73 + 7.03 X_t$$

The graph of the observed values and the trend line is displayed in Figure 1.

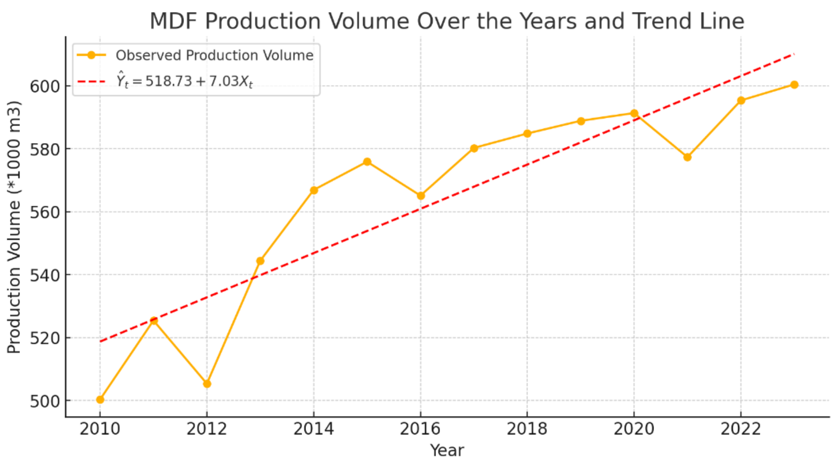

Figure 1: MDF production quantities by year and the trend line

In the graph, the time series data for MDF production quantities from 2010 to 2023 has been analyzed with observed values following a specific linear trend. The trend line modeling the change in production quantity is defined as:

$$\widehat{Y}_t = 518{,}73 + 7{,}03X_t \tag{5}$$

This linear model reflects a steady increase of approximately 7.03 units per year. Observed data indicate an overall increase in MDF production from 2010 onward, although deviations from the projected trend are noticeable in certain years. Such deviations suggest that production quantities may experience short-term fluctuations due to industrial, economic, or environmental factors impacting production.

Nevertheless, the general upward trend aligned with the linear model indicates growth in MDF production over the long term. This trend analysis can be instrumental in forecasting future production quantities, guiding industrial planning processes, and supporting strategic decisions regarding capacity expansion or resource allocation.

This trend analysis can be used to forecast production for future years, informing industrial planning processes and supporting strategic decisions such as capacity expansion or resource allocation.

Armutlulu (2008, pp. 292) proposed eight models that can be used to determine trends. These models, along with their respective linear transformations, are presented in Table 2.

Table 2: Eight models and their linear transformations for the trend determination

Model no	Model	Linear transformation
1	$Y_t = \alpha + bX$	$Y_t = \alpha + bX$
2	$Y_t = \alpha + \frac{b}{X}$	$Y_t = \alpha + b\frac{1}{X}$
3	$Y_t = \alpha X^b$	$LnY_t = Ln\alpha + bLnX$
4	$Y_t = \alpha + e^{bX}$	$LnY_t = Ln\alpha + bX$
5	$Y_t = \frac{1}{a + bX}$	$\frac{1}{Y_t} = a + bX$
6	$Y_t = \frac{X}{aX + b}$	$\frac{1}{Y_t} = a + b\frac{1}{X}$
7	$Y_t = \alpha + bLogX$	$Y_t = \alpha + bLogX$
8	$Y_t = e^{a+\frac{b}{X}}$	$LnY_t = a + b\frac{1}{X}$

Moving averages (MA)

Moving averages are a fundamental method for analyzing trends, seasonal fluctuations, and forecasting within time series data. According to Hyndman (2011), a moving average is a mathematical convolution created by taking the average of sequential values, which reduces short term variations, thereby clarifying the underlying trend in a series. This process diminishes the impact of transient fluctuations, allowing long-term patterns to become more prominent.

Moving averages are primarily employed in two forms: two-sided moving averages and one-sided moving averages. While two-sided moving averages are calculated using both past and present observations in the series, one-sided moving averages consider only past values. Hyndman (2011) notes that two-sided moving averages are particularly effective for highlighting underlying trends and seasonal components in time series data, whereas one-sided moving averages are more useful for short-term forecasting.

Additionally, there are centered and weighted versions of moving averages. Centered moving averages are particularly useful in removing seasonal fluctuations in data with strong seasonal components by averaging observations symmetrically around a given point. This approach provides a more balanced representation of the central trend. Weighted moving averages, on the other hand, assign varying weights to data points, which can result in a smoother representation of long-term trends.

In a time series with n elements, a s-period moving average is calculated by first taking the arithmetic mean of the initial s values. Then, the series progresses by one period, and the arithmetic mean of the next k values, starting from the second value up to the (s+1)-th value is calculated. This process continues throughout the series, smoothing out fluctuations and highlighting trends by averaging over consecutive subsets of data points.

Here, the first moving average value is positioned to align with the $\frac{(s+1)}{2}$-th value in the series. Similarly, the second moving average value is placed next to the subsequent data point in the series. This alignment continues throughout the series, ensuring that each moving average value corresponds to the central position of the k-period window it represents, providing a balanced view of the trend over time.

Revisiting the dataset in Example 1, Table 3 below represents the 3 period and 5 period MA fort his dataset.

Table 3: Calculation of 3 period and 5 period MA

Year	X	Production quantities (*1000 m^3)	3 MA	5 MA
2010	0	500.34		
2011	1	525.44	510.38	
2012	2	505.37	525.08	528.50
2013	3	544.44	538.91	543.61
2014	4	566.91	562.41	551.54
2015	5	575.88	569.30	566.52
2016	6	565.11	573.75	574.60
2017	7	580.25	576.73	578.99
2018	8	584.84	584.66	582.08
2019	9	588.88	588.35	584.53
2020	10	591.34	585.86	587.55
2021	11	577.35	588.01	590.67
2022	12	595.34	591.04	
2023	13	600.44		

As observed in Table 3 above, in the 3-period moving average, one period is lost from both the beginning and end of the series. In the 5-period moving average, two periods are lost from each end, creating new series based on the averages. Generalizing this, when the moving average is 3 or greater and consists of an odd number s, then $(s-1)/2$ periods are lost from both the start and end of the series, totaling $(s-1)$ periods lost overall.

If the moving average count, denoted as the value s, is an even number, as in the case of monthly data (e.g., 12-period MA) or quarterly data (e.g., 4-period MA), the moving average is computed in two stages. In the first stage, an s-period moving average is calculated. The first computed average is placed against the $(s+1)/2$-th period. However, this value is not an integer. Therefore, in the second stage, a double MA is taken to perform centering, with the first value positioned opposite the $(s+2)/2$-th period.

Seasonal fluctuations in time series

In time series analysis, seasonal fluctuations refer to regular, recurring patterns that occur over specific intervals within the data set, often influenced by seasonal or periodic factors such as months, quarters, or years (Hyndman and Athanasopoulos, 2018). These fluctuations represent changes in the time series that recur consistently due to influences tied to calendar-based cycles, including weather, holidays, and other cyclical factors. Detecting and adjusting for seasonal fluctuations is essential for accurately understanding the underlying trends and making reliable predictions in time series data.

Seasonal adjustments in time series analysis typically involve decomposing the data to isolate and remove these periodic effects, ensuring that the trend and cyclical components of the data are not obscured.

In identifying seasonal fluctuations within series, 12-period moving averages are typically calculated for monthly data, while 4-period moving averages are used for quarterly data. For instance, the 12-period moving averages calculated for monthly data correspond to the $\frac{12+1}{2} = 6.5$-th value. Consequently, as noted in the previous section, a double moving average is performed to achieve centering. As a result of this centering process, the first moving average value aligns with the seventh month.

Following this process, the ratio-to-moving average method is applied. By dividing each observed monthly value by the corresponding moving average value and then multiplying by 100, the percentage ratios to the moving average are obtained. This calculation provides the relative percentage deviation of each data point from its moving average, which aids in identifying and analyzing seasonal patterns within the time series.

Let us now revisit the data set for the annual MDF production quantities provided in Example 1. This time, however, we will evaluate this data set on a monthly basis. By converting the data to a monthly frequency, we can apply monthly moving averages and ratio-to-moving average methods, which will allow for a more granular analysis of potential seasonal fluctuations and trends over shorter intervals within the year.

Table 4: Calculation of the seasonal index for MDF production volume (*1000 m^3) using the Ratio to Moving Average Method

Year	Month	Y_t (*1000m^3)	2 MA	MA	Year	Month	Y_t (*1000m^3)	2 MA	MA
2010	Jan	2.48			2011	Jan	5.11		2.94
								3.01	
	Feb	3.58				Feb	5.85		3.09
								3.16	
	March	3.40				March	6.23		3.24
								3.32	
	Apr	3.64				Apr	5.69		3.41
								3.49	
	May	2.41				May	4.80		3.58
								3.67	
	June	2.01				June	4.06		3.73
			1.69					3.79	
	July	1.28		1.80		July	2.85		3.72
			1.91					3.65	
	Aug	9.62		2.01		Aug	1.42		3.59
			2.10					3.53	
	Sept	9.64		2.22		Sept	1.61		3.45

*Table 4: Calculation of the seasonal index for MDF production volume (*1000 m^3) using the Ratio to Moving Average Method*

Year	***Month***	***Y_t (*1000m^3)***	***2 MA***	***MA***	***Year***	***Month***	***Y_t (*1000m^3)***	***2 MA***	***MA***
			2.33					3.37	
	Oct	9.59		2.42		Oct	1.53		3.31
			2.50					3.25	
	Nov	0.65		2.61		Nov	2.89		3.17
			2.70					3.09	
	Dec	2.03		2.79		Dec	3.39		3.00
			2.87					2.91	
2022	Jan	1.36		9.14	**2023**	Jan	0.63		9.60
			9.20					9.61	
	Feb	2.09		9.25		Feb	1.69		9.64
			9.31					9.68	
	March	2.38		9.36		March	2.74		9.73
			9.41					9.79	
	Apr	2.14		9.42		Apr	2.79		9.85
			9.44					9.91	
	May	1.27		9.45		May	1.28		9.97
			9.47					0.03	

*Table 4: Calculation of the seasonal index for MDF production volume (*1000 m^3) using the Ratio to Moving Average Method*

Year	***Month***	***Y_t (*1000m³)***	***2 MA***	***MA***	***Year***	***Month***	***Y_t (*1000m³)***	***2 MA***	***MA***
	June	9.47		9.54		June	9.94		0.03
			9.61					0.04	
	July	8.31		9.58		July	8.39		
			9.55						
	Aug	6.83		9.53		Aug	7.68		
			9.52						
	Sept	6.97		9.53		Sept	8.38		
			9.55						
	Oct	6.77		9.55		Oct	8.22		
			9.56						
	Nov	7.68		9.56		Nov	9.07		
			9.56						
	Dec	0.05		9.58		Dec	0.14		
			9.60						

Table 5 has been created by reorganizing the centered 2-month moving average values in the last column of Table 4. In this new table, the seasonal index is obtained by calculating the arithmetic average for each month. However, while calculating the arithmetic average for each month, the outliers for that specific month were excluded from the calculation.

Table5: The calculation of the seasonal index based on the percentage ratios to the 12-month moving average

Month	2010	2011	.	.	.	2022	2023	Monthly Average Index	Adjusted Index
Jan		42.94	.	.	.	49.25	49.64	47.34	100.03
Feb		43.09	.	.	.	49.36	49.73	47.40	100.16
March		43.24	.	.	.	49.42	49.85	47.47	100.30
Apr		43.41	.	.	.	49.45	49.97	47.53	100.44
May		43.58	.	.	.	49.54	50.03	47.61	100.60
June		43.73	.	.	.	49.58		47.49	100.35
July	41.80	43.72	.	.	.	49.53		47.06	99.44
Aug	42.01	43.59	.	.	.	49.53		47.10	99.54
Sept	42.22	43.45	.	.	.	49.55		47.15	99.63
Oct	42.42	43.31	.	.	.	49.56		47.20	99.74
Nov	42.60	43.17	.	.	.	49.58		47.25	99.84
Dec	42.79	43.00	.	.	.	49.60		47.29	99.93

Since the index represents a percentage value, the total for the 12 months must equal 1200. Therefore, by performing a secondary adjustment in which we multiply the average of each month by the ratio $\left(\frac{1200}{Total\ Monthly\ Average\ Index}\right)$, we obtain the adjusted seasonal index.

Seasonal adjustment of time series

In the literature, various methods have been developed to seasonally adjust series by removing seasonal effects. These include

STL-Based Seasonal Adjustment with RegARIMA (Ollech, 2021), Bayesian Seasonal Adjustment (Holan and McElroy, 2012), State-Space Modelling (Harvey et al., 1997), and RSVD (Regularized Singular Value Decomposition) for Seasonal Decomposition (Lin et al., 2020). However, in this study, we will employ a more traditional approach, using the Ratio-to-Moving-Average method or the Seasonal Factor method to remove seasonal effects from the series.

In this method, the observed values are first divided by the adjusted index data obtained in Table 5 and then multiplied by 100 (*Equation 6*), effectively removing the seasonal effects from the data.

$$SA_t = \left(\frac{Y_t}{S_t}\right) * 100 \qquad (6)$$

where:

SA_t, represents seasonally adjusted value at time t, Y_t represents original series value at time t, and S_t shows seasonal factor at time t.

The new data, adjusted for seasonal fluctuations using the formula in Equation 6, is presented in Table 6 below.

Table 6: Monthly MDF production values adjusted for seasonal effects

Month	2010	2011	.	.	.	2022	2023
Jan	42.47	45.10	.	.	.	51.35	50.62
Feb	43.51	45.78	.	.	.	52.01	51.61
March	43.27	46.09	.	.	.	52.22	52.58
Apr	43.45	45.49	.	.	.	51.91	52.06
May	42.16	44.53	.	.	.	50.97	50.98
June	41.87	43.91	.	.	.	49.30	49.77
July	41.51	43.09	.	.	.	48.58	48.66
Aug	39.80	41.61	.	.	.	47.05	47.90
Sept	39.79	41.76	.	.	.	47.14	48.54
Oct	39.69	41.64	.	.	.	46.89	48.35
Nov	40.71	42.96	.	.	.	47.75	49.15
Dec	42.06	43.42	.	.	.	50.08	50.17

The graph illustrating the seasonal fluctuations in the original observation values is presented in Figure 2. Additionally, the graph for series adjusted for seasonal fluctuations is provided in Figure 4. In addition, Figure 5 shows the comparision of original and seasonally adjusted data.

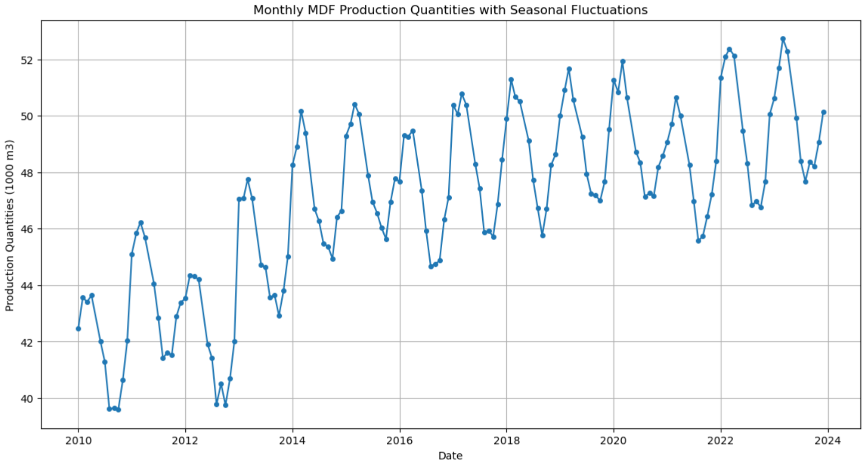

Figure 2: Seasonal fluctuation graph of observation values Y_t

If we compile the production amounts from the same month across different years to calculate an average value for each month, this approach allows us to establish a general monthly production average over the years, thereby enabling us to observe seasonal patterns. The resulting graph, as shown in Figure 3, illustrates these monthly averages.

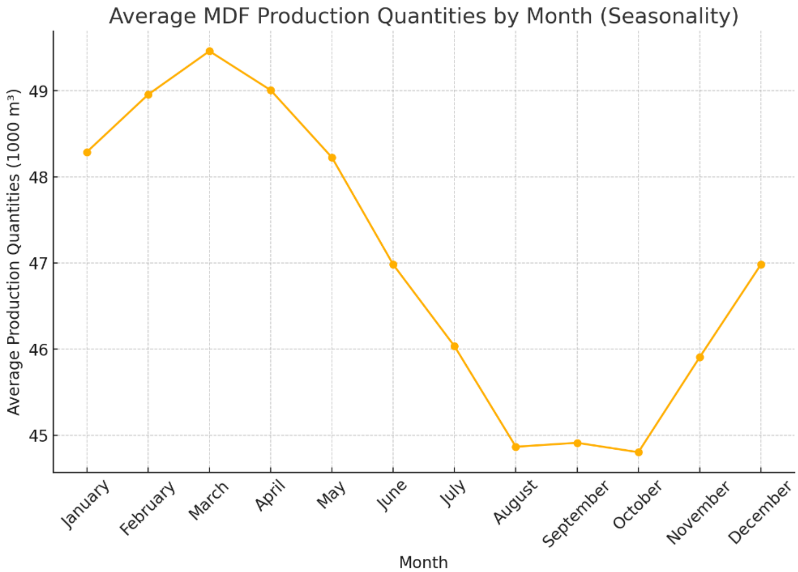

Figure 3: Average MDF production quantities by month

Since Figure 3 represents the overall monthly averages across years, it reveals the seasonal trend and allows us to observe recurring patterns at the monthly level.

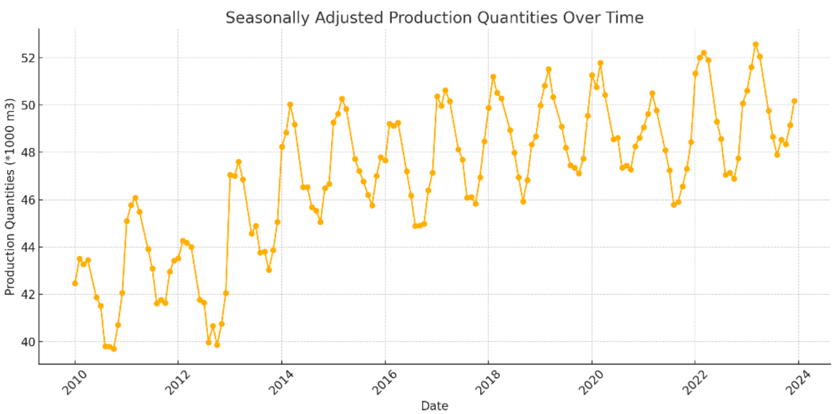

Figure 4: Seasonally adjusted production quatities over time

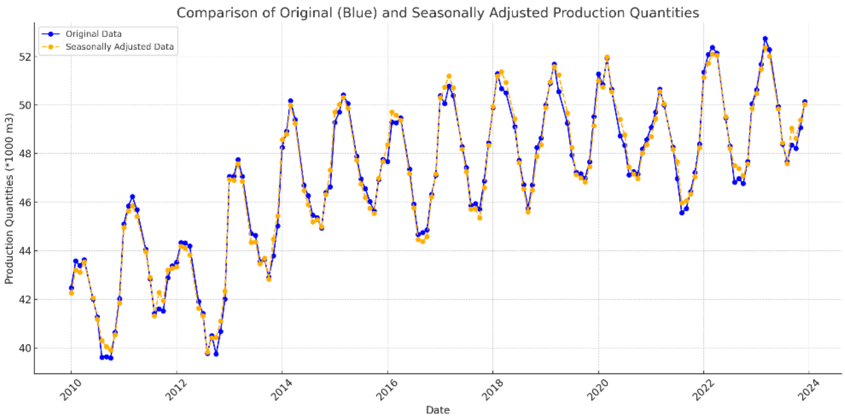

Figure 5: Comparision of original and seasonally adjusted data

The graph shows the deseasonalized monthly production amounts is presented in Figure 6. By removing the effect of the seasonal component in this graph, we can now observe the production amounts' fluctuations more clearly and irregularly. This adjustment facilitates a more straightforward analysis of cyclical or irregular variations.

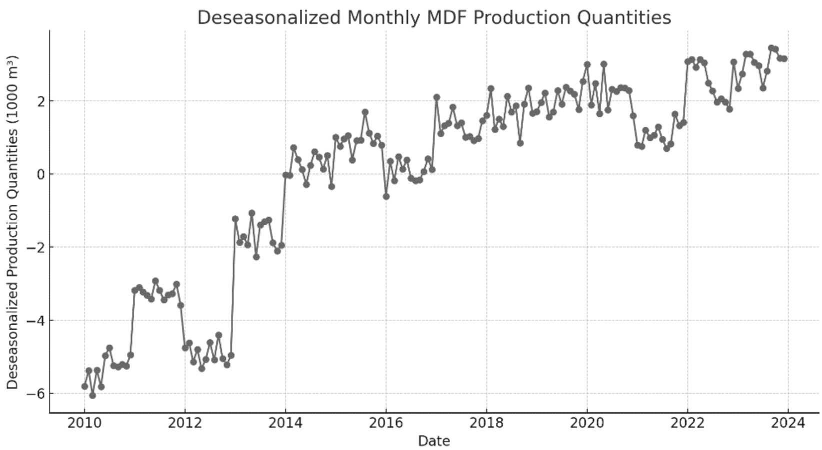

Figure 6: Deseasonalized monthly MDF production quantities

Cyclical and irregular fluctuations of time series

We can examine cyclical fluctuations using the formula $Y_t = T + S + C$. Here, cyclical fluctuations can be calculated as $C = Y_t - (T + S)$. In other words, by removing the trend component and seasonal fluctuations from our observed values, we can isolate and analyze cyclical fluctuations more effectively.

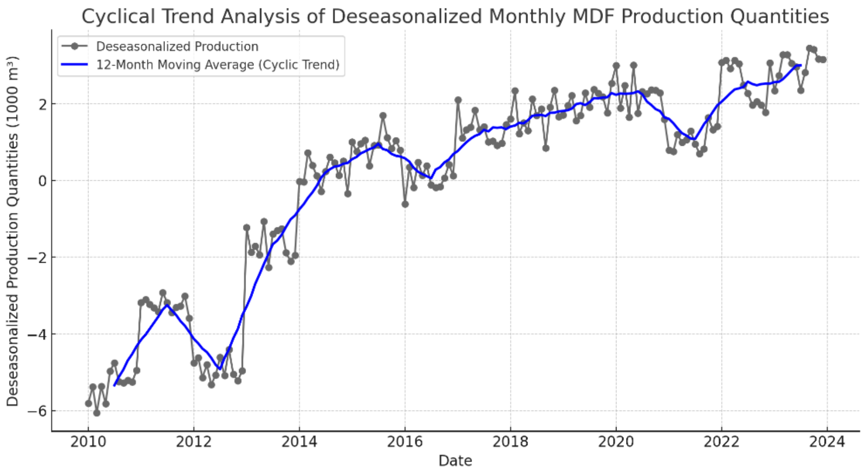

Figure 7: Cyclical trend analysis of deseasonalized monthly MDF production quantities

In the Figure 7, we can observe the long-term trend of seasonally adjusted production amounts showing cyclical fluctuations (blue line). This blue line has been generated using a 12-month moving average, which clarifies cyclical trends. Upon closer examination of the graph, certain years reveal upward or downward trends in production, indicating potential cyclical fluctuations. This approach allows for observing more extended and irregular variations in production levels.

To examine irregular fluctuations (I component), we need to remove the trend (T), seasonal (S), and cyclical (C) components from the current time series. Thus, the irregular fluctuation can be calculated as $I = Y_t - (T + S + C)$.

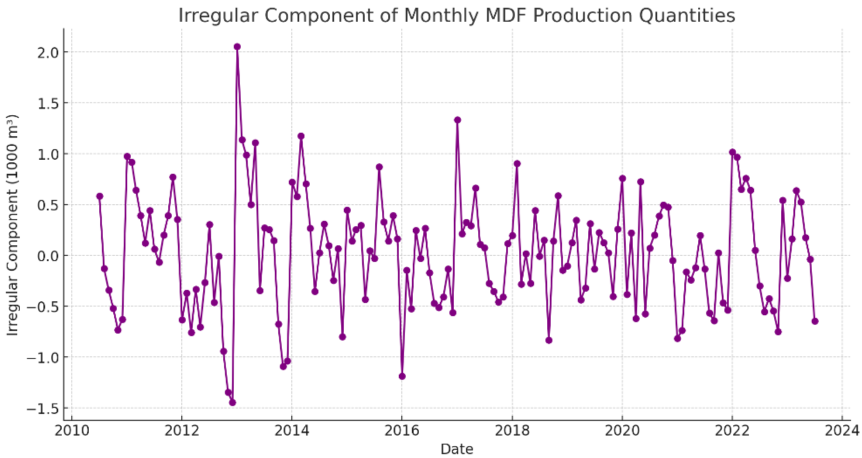

Figure 8: Irregular component of monthly MDF production quantities

In Figure 8, it can be observed the irregular component. This component has been adjusted to remove trend, seasonal, and cyclical effects, representing the remaining random fluctuations. The irregular component typically reflects changes caused by unexpected events, lacking any specific pattern or continuity. As this component does not exhibit a discernible pattern, it primarily reflects the influence of unpredictable, random factors.

Conclusion

In this chapter, a time series analysis was conducted on a sample dataset from the wood products sector, with a particular focus on the MDF industry, to examine sectoral trends, seasonal variations, and the cyclical and irregular components in detail. The primary objective of this study is to decompose the different components of the time series to gain a deeper understanding of production dynamics within the Forest Products Industry, develop forecasts concerning sectoral fluctuations, and support strategic decision-making.

Initially, the trend analysis applied to annual data reveals a general upward trend over the years. This finding suggests that the wood products sector exhibits a long-term growth tendency, indicating sustainable development within the industry. Furthermore, while trend analysis confirms a steady increase in annual production volume, it also highlights the potential vulnerability of this trend to economic fluctuations.

To facilitate a more detailed examination, the analysis then shifts to monthly data to focus on seasonal fluctuations. However, no distinct and recurring seasonal pattern was observed in the data, indicating that there is no regularly increasing or decreasing production volume during specific periods within the sector. This suggests that seasonal factors do not exert a notable influence on production, or that other factors such as external environmental or economic conditions might have a more dominant impact.

The data, adjusted to exclude seasonal effects, enabled a more precise analysis of cyclical fluctuations. In this analysis, cyclical trends were clarified, and long term, broader period fluctuations within the sector were identified. Although these fluctuations lack a strict periodicity, they point to long-term shifts that could arise due to various economic cycles, changes in demand, or supply demand imbalances within the industry. An accurate understanding of these cycles plays an essential role in strategic planning and resource management.

Finally, after removing trend, seasonal, and cyclical components, the remaining irregular component reflects the impact of unexpected and unpredictable short-term variations in MDF production. The presence of this irregular component suggests that production volumes in the sector are subject to momentary fluctuations, which introduce risk factors that must be considered in the production process.

Overall, time series analysis provides significant insights into the production dynamics of the wood products sector, offering valuable findings to guide long-term strategic decision-making. The methodological approaches offered by time series analysis are valuable not only in assessing current production data but also in anticipating potential risks and opportunities within the sector. In conclusion, this study highlights the application potential of time series analysis on sectoral data, thereby contributing to future studies and researchers in the wood products industry.

References

Armutlulu, İ. H. (2008). Sayısal yöntemler I: işletmelerde uygulamalı istatistik. Alfa yayınları.

Harvey, A., Koopman, S. J., & Riani, M. (1997). The modeling and seasonal adjustment of weekly observations. *Journal of Business & Economic Statistics*, 15(3), 354-368.

He, S., Zhao, X., Wang, E. Q., Chen, G. S., Chen, P. Y., & Hu, L. (2023). Engineered wood: sustainable technologies and applications. *Annual Review of Materials Research*, 53(1), 195-223.

Holan, S., & McElroy, T. (2012). Bayesian seasonal adjustment of long memory time series. *Economic Time Series: Modeling and Seasonality*, 403-429.

Huang, H., Su, S., & Li, L. (2024). Advancing timber construction: historical growth, research frontiers, and time series forecasting. *Journal of Asian Architecture and Building Engineering*, 1-30.

Huda, K., Wibowo, C. E., & Gunawan, V. (2023, June). Implementation of time series forecasting with Box Jenkins ARIMA method on wood production of Indonesian forests. In AIP *Conference Proceedings (Vol. 2738, No. 1)*. AIP Publishing.

Hyndman, R. J. & Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts.

Hyndman, J. R. (2011). Moving averages. *In International Encyclopedia of Statistical Science*. Springer, 866—869.

Kożuch, A., Cywicka, D., & Adamowicz, K. (2023). A comparison of artificial neural network and time series models for timber price forecasting. *Forests*, 14(2), 177.

Lin, W., Huang, J. Z., & McElroy, T. (2020). Time series seasonal adjustment using regularized singular value decomposition. *Journal of Business & Economic Statistics*, 38(3), 487-501.

Ollech, D. (2021). Seasonal adjustment of daily time series. *Journal of Time Series Econometrics*, 13(2), 235-264.

Szulecka, J. (2019). Towards Sustainable Wood-Based Energy: Evaluation and Strategies for Mainstreaming Sustainability in the Sector. *Sustainability*, 11(2), 493.


BİDGE