From Nature to Data, from Data to the Future: Ecological Independence and Digital Forestry

Editors
Prof. Dr. Gonca Ece OZCAN
Prof. Dr. Korhan ENEZ

DUVAR

FROM NATURE TO DATA, FROM DATA TO THE FUTURE: ECOLOGICAL INDEPENDENCE AND DIGITAL FORESTRY

Editors

Prof. Dr. Gonca Ece ÖZCAN

Prof. Dr. Korhan ENEZ

From Nature to Data, from Data to the Future: Ecological Independence and Digital Forestry
Editors: Prof. Dr. Gonca Ece ÖZCAN, Prof. Dr. Korhan ENEZ

Editor in chief: Berkan Balpetek
Cover Page Design: Arş. Gör. Ece ÇİLLİ
Page Design: Duvar Design
Printing : November -2025
Publisher Certificate No: 49837
ISBN: 978-625-8734-38-6

853 Sokak No:13 P.10 Kemeraltı-Konak/İzmir
Tel: 0 232 484 88 68
www.duvaryayinlari.com
duvarkitabevi@gmail.com

PREFACE

Forests, essential components of ecosystems, facilitate biodiversity conservation, human life sustainability, and the carbon cycle. Maintaining forest ecosystems and ensuring their sustainability is a current environmental and social requirement. Staying updated on contemporary technological advancements promotes innovative research in forestry. Remote sensing, artificial intelligence analyses, geographic information systems, and data science are becoming increasingly essential in contemporary forestry practices.

"***From Nature to Data, from Data to the Future: Ecological Independence and Digital Forestry***" thoroughly investigates the combination of digital forestry methodologies with the achievement of ecological independence. The book's investigations integrate field-based forestry methods with contemporary technologies, including satellite data, artificial intelligence, and statistical modeling, offering evaluations of forestry research and practice.

The book covers various topics, such as how landscape features affect insect damage assessments, measuring forest biomass, evaluating growing stocks, using machine learning over time, designing forest roads, and making decisions. These methodologies offer significant insights for both intellectual researchers and practitioners.

This book attempts to clarify novel perspectives in forestry research by integrating ecological understanding with technological advancements. We hope that readers will find this trip exciting and informative, recognizing the significance of evolving from data to future forestry practices.

Prof. Dr. Gonca Ece ÖZCAN
Prof. Dr. Korhan ENEZ

TABLE OF CONTENTS

Chapter 1

Evaluation of Pityo*kteines curvidens* Damage Based on Topographic Factors

Enes KAYA[1], Gonca Ece ÖZCAN[2*], Fatih SİVRİKAYA[3]

Introduction

Many factors, particularly climate change (Lieutier et al., 2016), adversely impact forests. Numerous abiotic and biotic variables, particularly bark beetle infestations, raise forest mortality (Weed et al., 2013). Bark beetles (Coleoptera: Curculionidae: Scolytinae) are prevalent, and outbreaks of certain species can lead to significant tree mortality throughout extensive forest regions (Vega and Hofstetter 2015). In recent years, tree mortality attributed to bark beetles has increased due to climate change (Fettig et al., 2022), resulting in substantial economic losses (Hlásny et al., 2021). The detrimental impacts of bark beetles affect the future sustainability of ecosystems (Hlásny et al. 2019).

The fir engraver beetle, *Pityokteines curvidens* (Germar, 1823), is a prevalent pest of fir trees (Pennacchio et al., 2002) and one of four bark beetle species within the genus Pityokteines (Pfeffer, 1995). It causes damage on the cambium layer of the lower trunk (Pernek, 2005), resulting in tree drying out (Kanat and Laz, 2005). This beetle inflicts damage on Abies species, with the exception of northern Europe (Charas, 1962). *P. curvidens* is a secondary pest in the western Black Sea forests of Türkiye (Küçük, 2001). Nevertheless, when conditions favor the species' development, it can reproduce quickly and cause harm on healthy forests (Schwerdtfeger, 1981), with significant population surges occurring in arid regions (Serin et al., 2005). The species typically favors thick-barked fir trees (Kanat and Laz, 2005) in forests with southern aspects and under stress (Şimşek, 2005). P. curvidens exhibits one or two generations annually, contingent upon elevation and temperature (Chararas, 1962). This species is

[1] Forestry Engineering, 37150, Kastamonu, Türkiye eneskaya18440@gmail.com
[2] Kastamonu University, Faculty of Forestry, Department of Forestry Engineering, Kastamonu, Türkiye, Orcid:0000-0003-0141-1031, goncaece@kastamonu.edu.tr
[3] Kastamonu University, Faculty of Forestry, Department of Forestry Engineering, Kastamonu, Türkiye, Orcid:0000-0003-0860-6747, fsivrikaya@kastamonu.edu.tr

anticipated to exacerbate tree mortality in European forests due to climate change (Seidl et al. 2017).

Efficient management of forest pests is essential for sustaining forests and ecological equilibrium (Özcan and Tabak, 2021). Efforts to maintain low population densities of these species and control measures implemented to prevent or mitigate the development of epidemics (Wermelinger 2004) may prove inadequate. The prompt and precise assessment of pest population numbers (Podlaski and Borkowski, 2009), the surveillance of these species, and the identification of factors influencing their harm are crucial for effective management techniques (Özcan et al., 2022b).

This study explored the variation in damage caused by *Pityokteines curvidens*, a bark beetle species, in fir forests based on topographic characteristics including slope, aspect, and elevation.

Material and Methods

Study Area

This research was conducted in the Erenler Forest Planning Unit (FPU) in the Kastamonu Regional Directorate of Forestry. The Erenler FPU is situated between 569000-584000 east longitudes and 4612000-4634000 north latitudes in the UTM coordinate system (WGS 84 Datum, Zone 36). The entire area of the research region is 14,516.6 hectares, of which 65% (9,372.2 hectares) is designated as forests. 78% (7282.8 ha) of the forest area comprises productive forest, while 22% (2089.4 ha) consists of degraded forest. The predominant tree species in the research area include black pine (*Pinus nigra*), Scots pine (*Pinus sylvestris*), fir (*Abies nordmanniana*), beech (*Fagus orientalis*), and oak (Quercus ssp.). *Pityokteines curvidens* caused significant damage to fir stands in the research area from 2019 to 2023. The extent of fir stands affected by the beetle in the research area is 578.2 hectares.

Database Development

The digital stand map was acquired from the Kastamonu Regional Directorate of Forestry. The Kastamonu Regional Directorate of Forestry efficiently executes beetle damage mitigation efforts within its jurisdiction. In this context, it executes "Forest Pest Control Projects" based on the year. This research utilized data on *P. curvidens* damage from 2019 to 2023. The damaged fir stands were incorporated into the stand map utilizing ArcGIS Pro. The Digital Elevation Model (DEM) was acquired from the website https://earthexplorer.usgs.gov. Elevation, slope, and aspect layers were generated utilizing DEM data. All

Geographic Information System (GIS) applications were conducted using ArcGIS Pro software.

The elevation, slope, and aspect layers were overlaid with the beetle-damaged stand layer to evaluate the extent of *P. curvidens* damage in relation to topographic parameters. The slope of the damaged stands was classified into four categories: less than 10%, 10-20%, 20-30%, and greater than 30%; aspect was divided into two categories: sunny and shady; and elevation was segmented into four categories: less than 1000 m, 1000-1250 m, 1250-1500 m, and greater than 1500 m.

Results

When assessing *P. curvidens* damage with respect to slope, 7.5% (253.2 ha) of the total area consisted of stands with a slope below 10%, 22.4% (759.8 ha) comprised stands with a slope of 10-20%, 26.9% (913.9 ha) included stands with a slope of 20-30%, and 43.2% (1463.7 ha) encompassed stands with a slope exceeding 30%. When the stands with beetle damage were evaluated, 5.7% (33.1 ha) of the damage occurred in stands with a slope of less than 10%, 19.3% (111.6 ha) in stands with a slope of 10-20%, 27.6% (159.3 ha) in stands with a slope of 20-30%, and 47.4% (274.2 ha) in stands with a slope of more than 30% (Table 1, Figure 1).

Table 1. Distribution of the study area and *P. curvidens* damage according to slope

Class (%)	Study Area		Damaged stands	
	ha	%	ha	%
<10	253.2	7.5	33.1	5.7
10-20	759.8	22.4	111.6	19.3
20-30	913.9	26.9	159.3	27.6
>30	1463.7	43.2	274.2	47.4
Total	3390.6	100	578.2	100

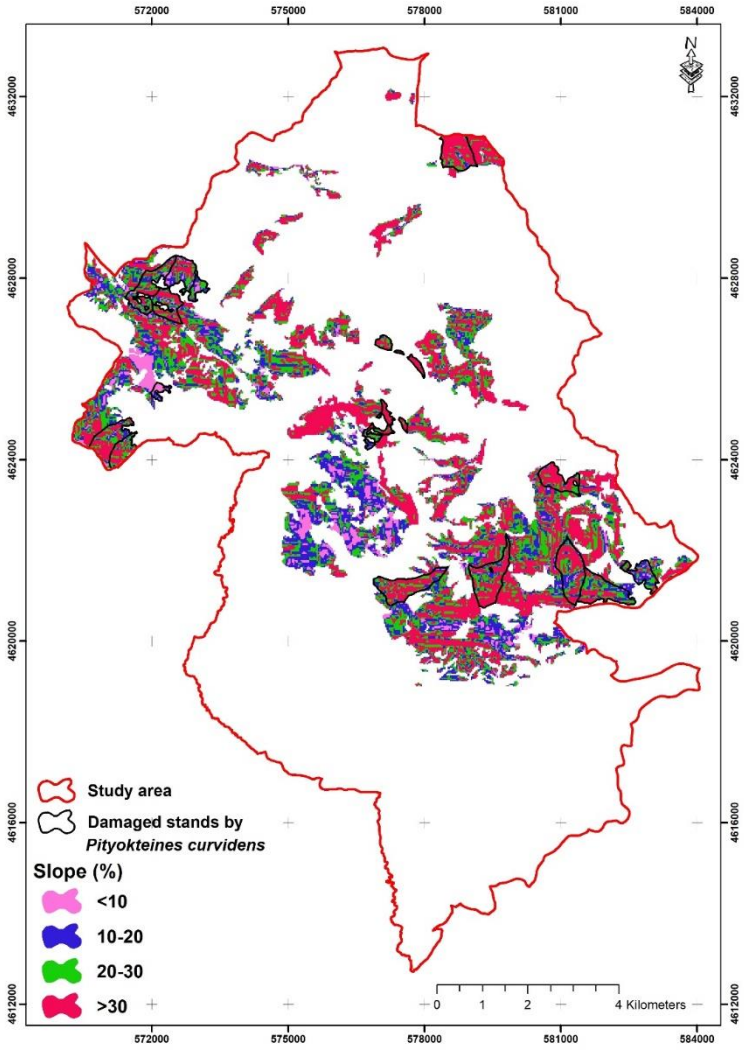

Figure 1. Distribution of the study area and *P. curvidens* damage to stands in different slope groups

Upon assessing *P. curvidens* damage by aspect, 63% (2134.9 ha) of the overall area was classified as sunny stands, while 37% (1255.7 ha) was categorized as shady stands. Based on evaluation of the stands exhibiting beetle damage, 82% (473.9 ha) of the damage occurred in sunny stands, while 18% (104.3 ha) was observed in shady stands (Table 2, Figure 2).

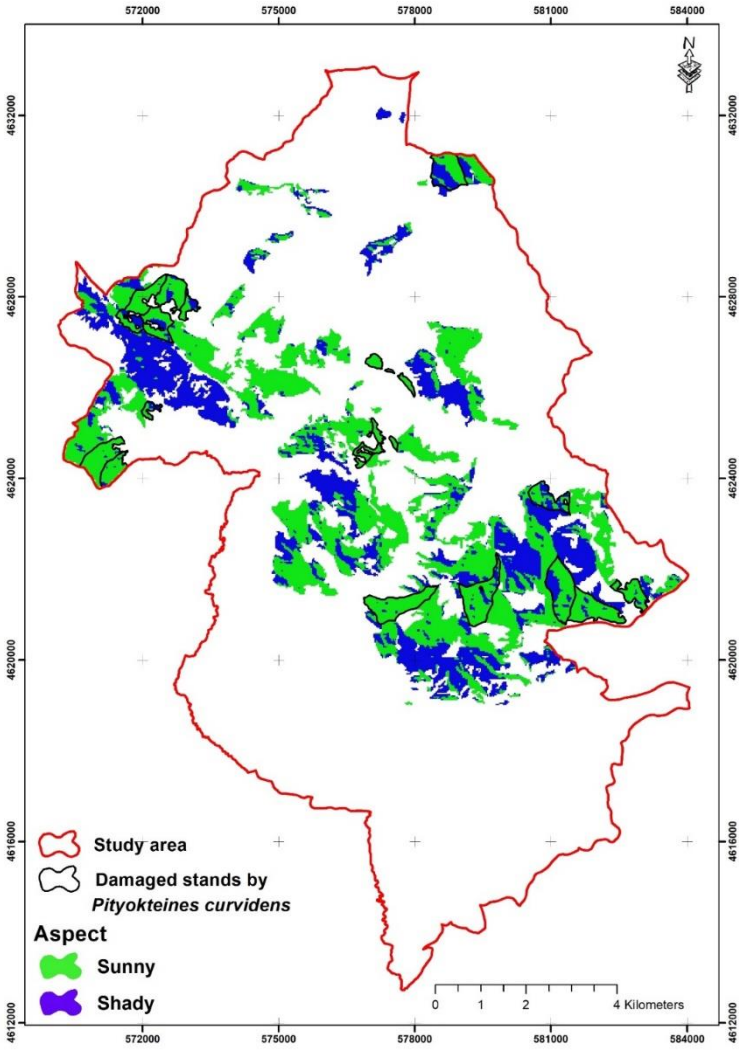

Figure 2. Distribution of the study area and *P. curvidens* damage to stands in different aspect groups

Table 2. Distribution of the study area and *P. curvidens* damage according to aspect

Class (%)	Study Area		Damaged stands	
	ha	%	ha	%
Sunny	2134.9	63.0	473.9	82.0
Shady	1255.7	37.0	104.3	18.0
Total	3390.6	100	578.2	100

In assessing *P. curvidens* damage by altitude, 1.5% (51.6 ha) of the total area was below 1000 m, 33.4% (1130.9 ha) was between 1000-1250 m, 59.5% (2017.1 ha) was between 1250-1500 m, and 5.6% (191.0 ha) was above 1500 m. Among the stands affected by beetles, 1.9% (11.2 ha) were located at elevations below 1000 m, 28.1% (162.2 ha) were situated between 1000-1250 m, 63.3% (366.2 ha) were found between 1250-1500 m, and 6.7% (38.6 ha) were at altitudes over 1500 m (Table 3, Figure 3).

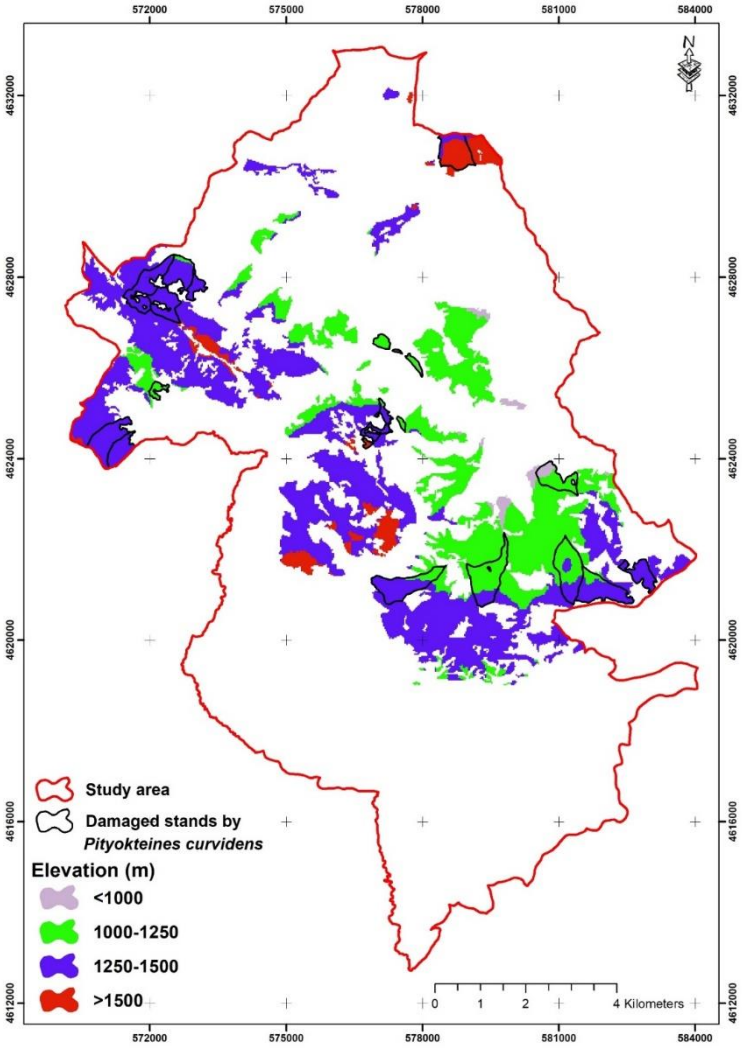

Figure 3. Distribution of the study area and *P. curvidens* damage to stands at different elevation groups

Table 3. Distribution of the study area and *P. curvidens* damage according to elevation

Class (%)	Study Area		Damaged stands	
	ha	%	ha	%
<1000	51.6	1.5	11.2	1.9
1000-1250	1130.9	33.4	162.2	28.1
1250-1500	2017.1	59.5	366.2	63.3
>1500	191.0	5.6	38.6	6.7
Total	3390.6	100	578.2	100

Discussion

A significant correlation exists between topographic parameters and bark beetle epidemics (Lausch et al., 2011). This research assessed the influences of slope, aspect, and elevation. Upon examining stands impacted by *Pityocteines curvidens* in relation to slope, it was ascertained that the proportion of injured stands escalated with increasing slope. The maximum damage occurred in areas with slopes exceeding 30%. The assessment of damage by elevation revealed that the percentage of damaged stands rose with elevation, except for those over 1500 m. It was concluded that fir stands were more susceptible to beetle attack. Numerous factors, including elevation, slope, aspect, temperature, precipitation, and stand age, directly or indirectly influence the damage and tree mortality inflicted by bark beetles in forests (Salinas Moreno et al. 2004; Samman and Logan 2000; Bale et al. 2002; Fora and Balog 2021).

A study on *Ips sexdentatus* identified that stand structure, crown closure, site index, stand age, slope, elevation, maximum temperature, precipitation, and solar radiation influence outbreak intensity (Özcan et al., 2022a). A study identifying forests susceptible to *Ips sexdentatus* employed analytical hierarchy process (AHP), frequency ratio (FR), and logistic regression (LR) models, revealing that stand structure, site index, crown closure, stand age, slope, elevation, maximum temperature, and solar radiation were significant influencing factors (Sivrikaya et al., 2022). Factors such as aspect, slope, elevation, stand type, and solar radiation are crucial in identifying forest regions vulnerable to *Ips typographus*, an aggressive species of bark beetle (Alkan Akıncı et al., 2022). A study employing the analytical hierarchy process (AHP) and MaxEnt models identified stand establishment, NDVI, and elevation as the primary variables predicting *P. curvidens* susceptibility (Sivrikaya et al., 2023). The correlation between bark beetle-induced tree mortality and climatic as well as topographic parameters should be assessed, considering regional conditions (Bałazy et al., 2019).

To maintain tree health and enhance ecosystem resilience, risk levels in forests must be assessed by evaluating various variables influencing bark beetle

infestations. A comprehensive evaluation and precise analysis of topographic data in risk assessments for the controlling of forest ecosystems will influence the efficacy of bark beetle management strategies.

References

Alkan Akıncı, H., Genç, Ç., Akıncı, H. (2022). Susceptibility assessment and mapping of *Ips typographus* (L.) (Coleoptera: Curculionidae) in oriental spruce forests in Artvin, Turkey. Journal of Applied Entomology, 00, 1–15.

Bałazy, R., Zasada, M., Ciesielski, M., Waraksa, P., Zawiła-Niedźwiecki, T. (2019). Forest dieback processes in the central European Mountains in the context of terrain topography and selected stand attributes. Forest Ecology and Management, 435:106–119.

Bale, J.S., Masters, G.J., Hodkinson, I.D., Awmack, C., Martijn Bezemer, T., Brown, V.K., Butterfield, J., Buse, A., Coulson, J.C., Farrar, J., Good, J.E.G., Harrington, R., Hartley, S., Jones, T.H., Lindroth, R.L., Press, M.C., Symrnioudis, I., Watt, A.D., Whittaker, J.B. (2002). Herbivory in global climate change research: direct effects of rising temperature on insect herbivores. Global Change Biology, 8: 1-16.

Chararas C (1962). Étude biologique des Scolytides des conifers. Paul Lechevalier, Paris

Fettig, C.J., Egan, J.M., Delb, H., Hilszczański, J., Kautz, M., Munson, A. S., Novak, J.T., Negron, J. F. (2022). Management tactics to reduce bark beetle impacts in North America and Europe under altered forest and climatic conditions. In *Bark beetle management, ecology, and climate change* (pp. 345-394). Academic Press.

Fora, C.G., Balog, A. (2021). The effects of the management strategies on spruce bark beetles populations (*Ips typographus* and *Pityogenes chalcographus*), in Apuseni Natural Park, Romania. Forests. 12, (6), 760.

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

Hlásny, T., Zimová, S., Merganičová, K., Štěpánek, P., Modlinger, R., & Turčáni, M. (2021). Devastating outbreak of bark beetles in the Czech Republic: Drivers, impacts, and management implications. Forest Ecology and Management, 490, 119075.

Kanat, M., Laz, B. (2005). Kahramanmaraş göknar ormanlarında *Pityokteines curvidens* (Germ.)'in feromon tuzaklarına yakalanma sonuçları. KSÜ Fen ve Mühendislik Dergisi, 8(2), 62-69.

Küçük, Ö., (2001). Batı Karadeniz ormanlarında göknar büyük kabukböceği (*Pityokteines curvidens* (Germ)) (Coleoptera: Scolytidae)'in son on yıldaki zararı ve mücadele çalışmalarının incelenmesi. Gazi Üniversitesi Kastamonu Orman Fakültesi Dergisi,1, 1, 53-62.

Lausch, A., Fahse, L., Heurich, M. (2011). Factors affecting the spatio-temporal dispersion of *Ips typographus* (L.) in Bavarian Forest National Park: A long-term quantitative landscape-level analysis. Forest Ecology and Management, 261: 233–45.

Lieutier, F., Mendel, Z., Faccoli, M. (2016). Bark beetles of mediterranean conifers. In: Paine, T., Lieutier, F. (Eds.) Insects and Diseases of Mediterranean Forest Systems, Springer, Cham, pp. 892.

Özcan, G.E., Sivrikaya, F., Sakici, O.E., Enez, K. (2022a). Determination of some factors leading to the infestation of *Ips sexdentatus* in crimean pine stands. Forest Ecology and Management, 519, 120316.

Özcan, G. E., Karacı, A., & Enez, K. (2022b). Application of deep neural networks in modeling the capture of *Ips sexdentatus* in pheromone trap. Environmental Engineering & Management Journal (EEMJ), 21(1): 105-116.

Özcan, G. E., Tabak, H. Ş. (2021). Evaluation of electronic pheromone trap capture conditions for *Ips sexdentatus* with climatic and temporal factors. Environmental Monitoring and Assessment, 193(10), 625.

Pennacchio, F., Gatti, E., Roversi, P.F. (2002). Le specie del genere Pityokteines Fuchs in Italia:Chiave identificativa e bionomia (Coleoptera, Scolytidae). Redia, 85, 229-256.

Pernek M (2005). Natural enemies of the fir bark beetles of the genus *Pityokteines* (Coleoptera, Scolytidae) in Croatia with special emphasis on pathogens. Dissertation. Faculty of Forestry, University of Zagreb, Zagreb.

Pfeffer, A. (1995). Zentral- und westpaläarktische Borken- und Kernkäfer. Pro Entomologia, Naturhistorisches Museum, Basel.

Podlaski, R., Borkowski, A. (2009). Estimating stem infestation density of *Pityokteines curvidens* (Germ.) on windfalls: a statistical approach. Journal of Pest Science, 82(4), 357-365.

Salinas-Moreno, Y., Mendoza, M.G., Barrios, M.A., Cisneros, R., Macias-Samano, J., Zuniga, G. (2004). Aerography of the genus Dendroctonus (Coleoptera: Curculionidae: Scolytinae) in Mexico. Journal of Biogeography, 31, 1163-1177.

Samman, S., Logan, J.A. (2000). Assessment and response to bark beetle outbreaks in the Rocky Mountain Area: report to congress from forest health protection, Washington Office, Forest Service, US Department of Agriculture. US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Schwerdtfeger, F. (1981). Waldkrankheiten. 4 Auflage. Paul Parey Verlag, Hamburg, Berlin

Seidl, R., Thom, D., Kautz, M., Martin-Benito., D, Peltoniemi, M., Vacchiano, G., Wild, J., Ascoli, D., Petr, M., Honkaniemi, J (2017). Forest disturbances under climate change. Nature Climate Change, 7: 395–402.

Serin, M., Erdem, M., Yüksel, B., Akbulut, S. (2005). Bolu ve Aladağ orman işletmesi göknar (*Abies bornmulleriana* Mattf.) ormanlarında etkin zarar yapan kabuk böceklerinin yaşam döngülerinin belirlenmesi ve bunlara karşı alınabilecek önlemlerin araştırılması. T.C. Çevre Ve Orman Bakanlığı, Batı Karadeniz Ormancılık Araştırma Müdürlüğü, Teknik Bülten No: 12.

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.

Sivrikaya, F., Özcan, G.E., Enez, K. (2023). Analytic hierarchy process - models, methods, concepts, and applications. Editors: Felice, D.E; Petrillo, A. Chapter 8: Predicting the susceptibility to *Pityokteines curvidens* using GIS with AHP and MaxEnt models in fir forests, IntechOpen, 153-168.

Şimşek, Z. (2005). Derbent (Ilgaz Dağı Milli Parkı) göknar orman alanlarında bulunan büyük göknar kabukböceği [*Pityokteines curvidens* (Germ.) (Coleoptera: Scolytıdae)]'nin zarar durumu ve uçuş periyodunun feromon tuzaklarla izlenmesi. Bartın Orman Fakültesi Dergisi, *7*(8), 18-26.

Vega, F.E, Hofstetter, R.W. (2015). Bark beetles: biology and ecology of native and invasive species. Academic Press.

Weed, A.S., Ayres, M.P., Hicke, J.A. (2013). Consequences of climate change for biotic disturbances. Ecological Monographs, 83:441–470.

Wermelinger, B. (2004). Ecology and management of the spruce bark beetle *Ips typographus*-a review of recent research. Forest Ecology and Management, 202:67–82.

Chapter 2

Estimation of Aboveground Biomass in Pure Cedar Stands Using Landsat-8 OLI Satellite Imagery: A Case Study from the Kaş Forest Enterprise Directorate

Fatih SARIÇAM[1], Sedat KELEŞ[2]

Introduction

Aboveground biomass (AGB) refers to the total living organic matter of plant organisms within an ecosystem, particularly trees, that remains above the soil surface. This concept encompasses the sum of tree components such as stems, branches, leaves, and bark. AGB plays a central role in monitoring both ecological and climatic processes; it is widely employed as a key parameter in studies concerning the understanding of the carbon cycle, the monitoring of atmospheric greenhouse gas concentrations, and the assessment of carbon sinks (Sun et al., 1980; Saraçoğlu, 1988; Dixon et al., 1994).

The increasing global demand for energy has made the use of renewable energy sources essential in order to reduce dependence on fossil fuels. In this context, forest biomass has gained growing importance both for sustainable energy production and for carbon balancing strategies. In countries rich in forest resources, such as Türkiye, accurately assessing the biomass potential obtainable from forests is of vital importance for the sustainable management of natural resources. Moreover, monitoring the structural changes that forests undergo over time requires the evaluation of biomass stocks with up-to-date data (Sun et al., 1980).

Traditional forest inventory methods rely on field-based measurements, in which individual tree attributes such as diameter at breast height, height, and species are recorded. Although reliable in terms of accuracy, these methods are limited in their applicability over large areas, as well as being costly and time-consuming. Consequently, remote sensing technologies and geographic

[1] General Directorate of Forestry, Ankara, Türkiye, Orcid:0009-0005-9244-0664, fatihsaricam@gmail.com
[2] Çankırı Karatekin University, Faculty of Forestry, Department of Forest Engineering, Çankırı, Türkiye, Orcid:0000-0002-2724-983X, dr.sedatkeles@gmail.com

information systems (GIS) have emerged as effective alternatives for the estimation and monitoring of forest biomass. Spectral band reflectance values, vegetation indices (e.g., NDVI, EVI), and texture data derived from satellite imagery provide indirect yet robust indicators of the physiological and structural characteristics of trees.

In Türkiye, research on estimating aboveground biomass using remote sensing techniques has increased significantly in recent years. For instance, Landsat and Sentinel satellite data have been applied to pure stands of Anatolian black pine (*Pinus nigra*) and Turkish red pine (*Pinus brutia*), where spectral band reflectance values, vegetation indices, and texture measures have been utilized for biomass estimation (Günlü et al., 2014; Turgut, 2019; Işık, 2024; Güverçin, 2022). In these studies, multiple regression analysis, deep learning and support vector machines approaches have been employed to improve model performance (Keleş et al., 2021). The results demonstrate that combining different satellite datasets with advanced modeling techniques enhances prediction accuracy.

Globally, a broader range of studies has been conducted across diverse climatic zones and forest types. In particular, high-resolution LiDAR data, as well as MODIS and QuickBird imagery, have yielded high levels of accuracy in biomass modeling (Baccini et al., 2008; Brice et al., 2013; Hirata et al., 2013; Dube and Mutanga, 2015). From tropical forests to boreal regions, these studies reveal that remote sensing techniques are effective tools for biomass estimation and are now widely adopted at a global scale (Zheng et al., 2004; Lu, 2005).

The objective of this study is to contribute to the rapid, cost-effective, and widespread monitoring of biomass in forest ecosystems through remote sensing technologies, while also providing a scientific foundation for sustainable forest management, climate change mitigation, and carbon emission reduction policies. To this end, aboveground biomass calculations were conducted using diameter at breast height data collected from field studies in selected sample plots within the Kaş Forest Enterprise Directorate, located in southern Türkiye, under the 2024 revised forest management plan. The biomass values obtained were then related to vegetation indices, spectral band reflectance data, and texture measures derived from Landsat 8 OLI imagery. Analyzing these relationships through statistical modeling techniques is expected to enable both the prediction of the spatial distribution of biomass and the generation of highly accurate biomass maps based on remote sensing data.

Material and Methods

Case Study Area

The study area was identified as the Kaş Forest Enterprise Directorate, located within the boundaries of the Antalya Regional Directorate of Forestry (Figure 1). Geographically, it lies between 36°39′06″–36°07′22″ N latitude and 29°44′17″–29°15′40″ E longitude. The average elevation of the study area is 1,034 m above sea level. The annual mean temperature is 19.6°C and an average annual precipitation is 782.9 mm. The total surface area of the study region is 180,335.70 ha, of which 57% (114,780.0 ha) is forested land and 43% (57,177.50 ha) is non-forested land. The tree species distributed across the area include cedar (*Cedrus libani*), black pine (*Pinus nigra*), Turkish red pine (*Pinus brutia*), valonia oak (*Quercus ithaburensis* subsp. *macrolepis*), holm oak (*Quercus ilex*), kermes oak (*Quercus coccifera*), juniper (*Juniperus* spp.), stone pine (*Pinus pinea*), and eucalyptus (*Eucalyptus* spp.) (OGM, 2024).

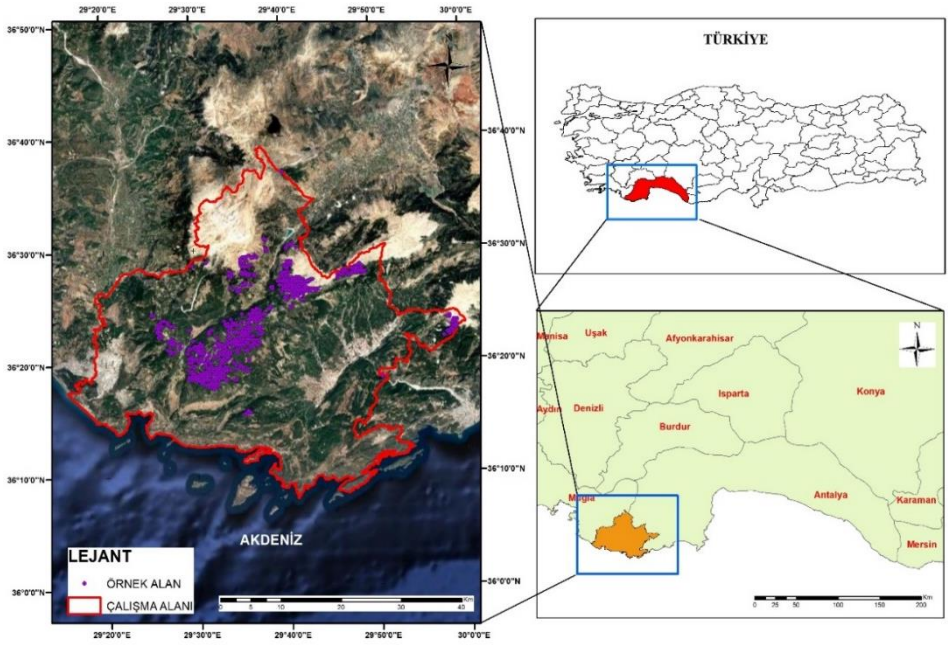

Figure 1. Location of the study area

Material

The materials used included the stand type map of the Kaş Forest Enterprise Directorate under the Antalya Regional Directorate of Forestry, a Landsat 8 OLI satellite image dated July 30, 2024, and inventory card data from 618 sample plots collected between June and August 2024. The dataset employed in this study consists of inventory card records gathered by the forest management team in 2024 from pure cedar (*Cedrus libani*) stands, as part of the revision of the forest management plan of the Kaş Forest Enterprise Directorate.

Method

Calculation of Aboveground Biomass

In 2024, forest inventory studies were conducted within the framework of the forest management plans of the Kaş Forest Enterprise Directorate. From the data obtained in these inventories, 618 pure cedar (*Cedrus libani*) stands were selected. For each sample plot, diameter at breast height (DBH, d1.3) values of trees recorded in the inventory cards were used, and the aboveground biomass (kg) of each tree was calculated using Equation (1) below. The biomass estimation was based on the model developed by Durkaya et al. (2013), as presented below. The total AGB of each sample plot was derived by summing the biomass of all trees within the plot. Considering the plot sizes (400 m², 600 m², and 800 m²), the values were converted to a per-hectare basis, and the aboveground biomass per hectare was calculated.

$$AGB = 37.21449 + (-8.08322 * d_{1.3}) + (0.644812 * d_{1.3}^2) \quad \ldots\ldots \quad (1)$$

Where:

AGB (ton): Aboveground biomass

$d_{1.3}$: Diameter at breast height (DBH) of the tree, obtained from inventory cards

Evaluations Regarding Landsat 8 OLI Satellite Imagery

Landsat 8 OLI satellite imagery was utilized in this study. The related satellite data were obtained free of charge from https://earthexplorer.usgs.gov. The spatial resolution of Landsat 8 OLI is 30 m, and Bands 2, 3, 4, 5, 6, and 7 were employed in the analysis. Some fundamental information regarding the Landsat 8 OLI satellite imagery is presented in Table 1.

Table 1. Specifications of Landsat 8 OLI bands

	Bands	Wavelength (nμ)	Spatial Resolution (m)
Landsat 8 OLI	Band2	0,45 – 0,51	30
	Band3	0,53 – 0,59	30
	Band4	0,67 – 0,67	30
	Band5	0,85 – 0,88	30
	Band6	1,57 – 1,65	30
	Band7	2,11 – 2,29	30

Calculation of Band Reflectance Values from Landsat 8 OLI Satellite Imagery

The Landsat 8 OLI imagery was subjected to preprocessing through atmospheric correction for subsequent analysis. Within the scope of this study, six bands (Band 2, Band 3, Band 4, Band 5, Band 6, and Band 7) were converted into reflectance images using QGIS software. Sampling areas were delineated on the Landsat 8 OLI bands, and the reflectance values corresponding to each sampling area were calculated. The band reflectance values of the sampling plots were derived by calculating the mean reflectance of pixels located within the buffer zones, whose radii were determined according to the specified plot sizes (11.28 m for 400 m², 13.82 m for 600 m², and 15.96 m for 800 m²). This process was implemented in ArcGIS 10.5 software using the ArcToolbox → Spatial Analyst Tools → Zonal → Zonal Statistics module.

Derivation of Vegetation Index Values Using Landsat 8 OLI Satellite Imagery

Based on the band reflectance values computed for the sampling plots, the following vegetation indices were derived: GNDVI (Green Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index), MID (Middle Infrared Wavelengths), NDVI (Normalized Difference Vegetation Index), MSI (Moisture Stress Index), MSR (Modified Simple Ratio), NDMI (Normalized Difference Moisture Index), NLI (Non-linear Index), NBR (Normalized Burn Ratio), PSSR (Pigment Specific Simple Ratio), EVI (Enhanced Vegetation Index), IPVI (Infrared Percentage Vegetation Index), NBR2 (Normalized Burn Ratio 2), SAVI (Soil Adjusted Vegetation Index), RVI (Ratio Vegetation Index), MSAVI (Modified Soil Adjusted Vegetation Index), EVI2 (Enhanced Vegetation Index 2), ARVI (Atmospherically Resistant Vegetation Index), DVI (Difference Vegetation Index), ND32 (Normalized Difference 32), MID57 (Middle Infrared Wavelengths 57), ND35 (Normalized Difference 53), VISI23 (Visible Bands), and ND73 (Normalized Difference 73).

Derivation of Texture Values from Landsat 8 OLI Satellite Imagery

Eight different texture features were utilized, which are defined as follows: variance (VAR), mean (M), contrast (CON), entropy (ENT), dissimilarity (DIS), second moment (SM), homogeneity (HOM), and correlation (COR). For each band of the satellite imagery, these texture features were calculated, and texture values for each sample plot were derived using window sizes of 3×3, 5×5, 7×7, and 9×9 pixels. These calculations were performed using ENVI software. Detailed information regarding the texture features obtained in this study is presented in Table 2.

Table 2. Abbreviations and symbols for certain texture features used in this study

Texture Features	Bands	Window Size			
		3x3	5x5	7x7	9x9
Contrast	Band2	Band2 3x3 COR	Band2 5x5 COR	Band2 7x7 COR	Band2 9x9 COR
Correlation	Band3	Band3 3x3 CON	Band3 5x5 CON	Band3 7x7 CON	Band3 9x9 CON
Dissimilarity	Band4	Band4 3x3 DIS	Band4 5x5 DIS	Band4 7x7 DIS	Band4 9x9 DIS
Entropy	Band5	Band5 3x3 HOM	Band5 5x5 HOM	Band5 7x7 HOM	Band5 9x9 HOM
Homogeneity	Band6	Band6 3x3 ENT	Band6 5x5 ENT	Band6 7x7 ENT	Band6 9x9 ENT
Mean	Band7	Band7 3x3 V	Band7 5x5 V	Band7 7x7 V	Band7 9x9 V
Second Moment					
Variance					

Statistical Analysis

In this thesis, stepwise multiple regression analysis was employed to model statistical relationships. The modeling process was carried out using the SPSS statistical software. Regression models were developed between independent variables derived from satellite imagery and the dependent variable (aboveground biomass). The structure of the regression model predicting the relationships between the independent variables (vegetation indices, texture, and band reflectance) and the dependent variable (aboveground biomass) is presented in Equation 2.

A total of 618 inventory plot data were used in this study. Of these, 80% (461 plots) were used for model development, while 20% (157 plots) were

reserved for model validation. Paired t-tests were applied to evaluate the performance of the models.

$$\text{Aboveground Biomass}=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\ldots+\beta_nX_n+\varepsilon \quad (2)$$

In this equation, β0,β1,β2,…,βn represent the regression coefficients, while X1,X2,X3,…,Xn denote the texture features, band reflectance values, and vegetation indices calculated for the sample plots from Landsat 8 OLI satellite imagery. The term ε represents the error in the model.

Using Landsat 8 OLI satellite data, statistically significant model options were determined by combining band reflectance values, texture features, and vegetation indices at a 95% confidence level.

Results

The relationships between the independent variables (band reflectance) calculated for the sample plots and the dependent variables (aboveground biomass) were examined using multiple regression analysis (Table 3). In the model including Landsat 8 OLI band reflectance variables, the coefficient of determination was found to be R^2_{adj}=0.186. The graphs related to the modeling of aboveground biomass based on band reflectance values are presented in Figure 2.

Table 3. Model results with the highest accuracy for aboveground biomass estimation based on band reflectance values

Independent Variables	Regression Coefficients	Standard Error	t-statistics	p
	-1532,878	231,516	-6,621	0,000
B7	-4005,718	420,612	-9,524	0,000
B2	9239,289	1315,720	7,022	0,000
B5	361,294	157,229	2,298	0,022
$R^2_{düz}$=0.186 Sy. x= 78,5729				
AGB (ton ha-1) = -1532,878 - (4005,718 x B7) + (9239,289 x B2) + (361,294 x B5)				

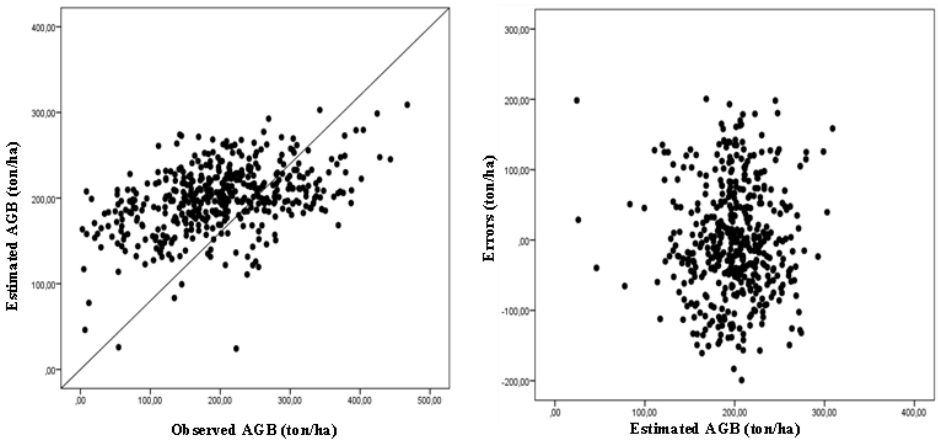

Figure 2. Model results of aboveground biomass based on band reflectance values

The model outcomes with the highest predictive accuracy, based on vegetation indices obtained from Landsat 8 OLI data, are shown in Table 4. Upon examining Table 4, the model including Landsat 8 OLI vegetation index variables was found to have a coefficient of determination of R^2_{adj}=0.181. Compared to the model based on band reflectance values, similar results were obtained for the model incorporating vegetation index variables. The graphs related to the modeling of aboveground biomass based on band vegetation indices are presented in Figure 3.

Table 4. Findings on vegetation index values

Independent Variables	Regression Coefficients	Standard Error	t-statistics	p
	-1741,833	274,132	-6,354	0,000
NBR	756,637	114,644	6,600	0,000
VIS123	2764,687	448,972	6,158	0,000
SARVI	-29,916	8,038	-3,722	0,000
ND32	-1750,603	773,026	-2,265	0,024
R^2_{adj}=0,181 Sy, x=77,5853				

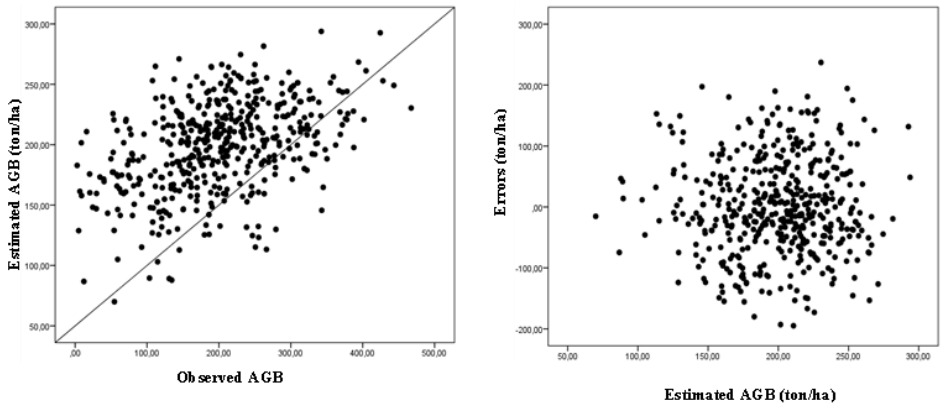

Figure 3. Model results of aboveground biomass based on vegetation indices

Upon examining Table 5, the model based on Landsat 8 OLI independent variables (texture) was found to have a coefficient of determination of R^2_{adj}=0.453. Compared to the models developed using band reflectance and vegetation index values, the model based on texture variables demonstrated a higher level of explanatory power. The graphs related to the modeling of aboveground biomass based on band texture values are presented in Figure 4.

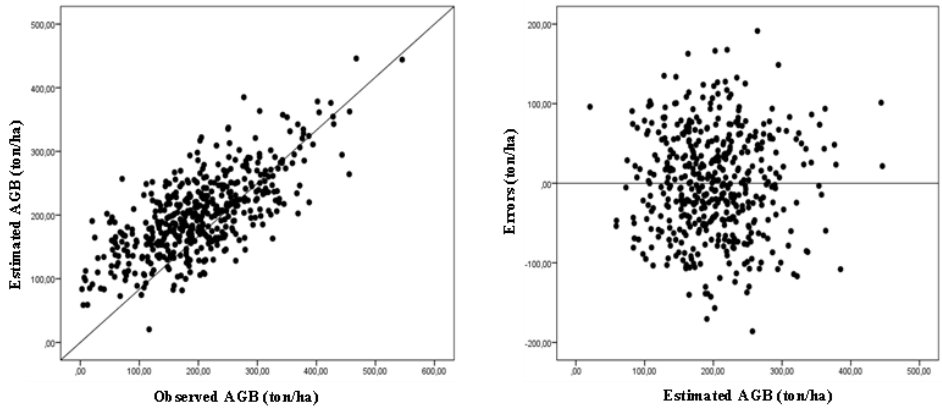

Figure 4. Model results of aboveground biomass based on texture values

Table 5. Results Related to Texture Features

Independent Variables	Regression Coefficients	Standard Error	t-statistics	p
(Constant)	-818,207	216,470	-3,780	,000
DİSSİMİLARİTY_BANT5_9x9_0	311,245	48,806	6,377	,000
SECOND MOMENT_BANT6_5x5_45	602,959	87,776	6,869	,000
CONTRAST_BANT6_5x5_90	11,189	3,769	2,969	,003
MEAN_BANT7_3x3_135	-55,630	7,396	-7,522	,000
MEAN_BANT5_3x3_135	10,891	3,935	2,768	,006
CONTRAST_BANT7_3x3_135	8,555	3,448	2,481	,013
SECOND MOMENT_BANT6_5x5_0	-171,892	91,397	-1,881	,041
ENTROPY_BANT7_5x5_135	-41,160	10,853	-3,793	,000
SECOND MOMENT_BANT5_3x3_135	-111,517	22,579	-4,939	,000
HOMEGENEİTY_BANT5_7x7_0	403,770	110,603	3,651	,000
CONTRAST_BANT2_3x3_0	-93,653	18,451	-5,076	,000
CONTRAST_BANT3_3x3_45	39,933	9,050	4,412	,000
ENTROPY_BANT6_3x3_45	57,301	14,383	3,984	,000
SECOND MOMENT_BANT6_7x7_0	-609,547	107,459	-5,672	,000
MEAN_BANT3_5x5_0	66,555	13,570	4,904	,000
VARİENCE_BANT7_3x3_135	23,711	7,260	3,266	,001
ENTROPY_BANT4_9x9_135	62,047	12,114	5,122	,000
VARİENCE_BANT3_5x5_0	-64,287	19,624	-3,276	,001
ENTROPY_BANT6_9x9_0	-101,327	21,067	-4,810	,000
SECOND MOMENT_BANT4_5x5_0	76,274	22,202	3,435	,001
CORRELATİON_BANT5_9x9_0	97,142	27,513	3,531	,000
CONTRAST_BANT6_7x7_0	12,265	5,103	2,404	,017
CORRELATİON_BANT6_9x9_45	-64,236	20,608	-3,117	,002
CORRELATİON_BANT5_5x5_0	-36,342	16,366	-2,221	,027
ENTROPY_BANT6_3x3_135	-33,199	10,317	-3,218	,001
CORRELATİON_BANT2_5x5_135	-24,963	8,840	-2,824	,005
CORRELATİON_BANT3_3x3_135	-24,898	7,475	-3,331	,001
CORRELATİON_BANT3_3x3_90	21,773	8,299	2,623	,009
CORRELATİON_BANT4_9x9_0	46,584	20,252	2,300	,022
HOMEGENEİTY_BANT2_9x9_0	415,695	131,780	3,154	,002
DİSSİMİLARİTY_BANT2_5x5_0	124,017	54,323	2,283	,023

R^2_{adj}=0,453 Sy,x=66,9189

The most successful model developed for predicting aboveground biomass, which included texture variables, was tested using an independent dataset. A paired t-test was conducted to evaluate the model's suitability for the sampled population by comparing predicted and observed AGB values. The analysis

indicated t = 0.093 and p = 0.926, showing no statistically significant difference between predicted and observed values at $\alpha = 0.05$ ($p > 0.05$). These results suggest that the regression model incorporating texture variables is well-suited for the pure cedar stands used in this study.

Discussion

The relationships between aboveground biomass and various independent variables derived from Landsat 8 OLI imagery were evaluated using multiple regression analysis. Models based on band reflectance, texture variables, and vegetation indices were compared in terms of their explanatory power. The model incorporating band reflectance values yielded a coefficient of determination of R^2=0.186, while the model based on vegetation index values produced a similar R^2=0.181 (Table 4). In contrast, the model developed using texture variables achieved a substantially higher coefficient of determination, R^2=0.453 (Table 5), indicating that texture features provide a stronger predictive capability for aboveground biomass compared to band reflectance or vegetation index values. These results suggest that incorporating texture information from Landsat 8 OLI imagery enhances the accuracy of biomass estimation models.

Numerous studies in the literature have focused on the estimation of AGB using various RS data. Işık (2024) employed Landsat 8 OLI imagery to estimate aboveground stand carbon in pure black pine forests, reporting determination coefficients of R^2 = 0.593 for band reflectance, R^2 = 0.635 for vegetation indices, and R^2 = 0.581 for texture features. Güverçin (2022) estimated AGB using allometric equations and modeled the relationship between Landsat 8 OLI imagery and Sentinel-1 AGB data via multiple regression, achieving a maximum R^2 = 0.509 for texture, band reflectance, and vegetation indices. Keleş et al. (2021) assessed AGB estimation using combined Sentinel-1 and Sentinel-2 data with regression, deep learning, and support vector machine (SVM) models. Results showed that data integration significantly improved model performance, with R^2 = 0.877 for SVM, R^2 = 0.85 for deep learning, and R^2 = 0.748 for regression models. Turgut (2019) applied multiple regression analysis, obtaining R^2 values of 0.445, 0.387, and 0.552 for vegetation indices, band reflectance, and texture, respectively.

Dube and Mutanga (2015) estimated AGB for pine and eucalyptus using Landsat 7 ETM+ and Landsat 8 imagery, with band reflectance R^2 values ranging from 0.21–0.43 (Landsat 7) and 0.30–0.47 (Landsat 8), and vegetation index R^2 values from 0.33–0.47 (Landsat 7) and 0.36–0.58 (Landsat 8). Günlü et al. (2014) estimated AGB in black pine forests using Landsat TM imagery, reporting R^2 = 0.465 for band reflectance and R^2 = 0.606 for vegetation indices.

Brice et al. (2013) utilized LiDAR and high-resolution spatial data to model forest volume, tree height, and biomass, achieving $R^2 = 0.76$ for tree height, $R^2 = 0.94$ for volume, and $R^2 = 0.92$ for biomass. Hirata et al. (2013) examined allometric relationships between forest canopy closure and AGB using high-resolution QuickBird imagery, yielding $R^2 = 0.65$. Baccini et al. (2008) applied MODIS medium-resolution data to estimate AGB in tropical African forests, achieving $R^2 = 0.90$ between predicted AGB and field-measured heights, with cross-validation confirming 82% of AGB variance explained.

Conclusion

The current research focused on examining the relationships between aboveground biomass and texture features, vegetation indices, and band reflectance derived from Landsat 8 OLI imagery using multiple regression analysis. Three separate models were developed based on these independent variables, with the model incorporating texture features achieving the highest predictive performance. Models based on vegetation indices and band reflectance yielded similar but lower accuracies.

A review of the literature indicates that studies using Landsat imagery generally achieve moderate model performance, with prediction accuracy often insufficient. Factors contributing to lower model performance may include image acquisition timing, errors in preprocessing, terrain characteristics, low spatial resolution, and the modeling techniques applied. Future studies could enhance predictive accuracy by employing higher-resolution satellite imagery and advanced modeling approaches. Specifically, integrating variables derived from LiDAR data, high-resolution passive imagery, long-wavelength active imagery (e.g., P- and L-bands), and advanced techniques such as artificial neural networks or deep learning may notably increase the accuracy of aboveground biomass estimation models.

Acknowledgements

This book chapter is derived from the master's thesis entitled *"Estimation of Above-Ground Biomass Using Landsat 8 OLI Satellite Imagery in Pure Cedar Stands (Case Study of Kaş Forest Enterprise Directorate)"*, conducted by Fatih Sarıçam in 2024 at the Institute of Science, Çankırı Karatekin University, under the supervision of Prof. Dr. Sedat Keleş. We would like to express our gratitude to the General Directorate of Forestry, Department of Forest Management and Planning, for providing the data used in this study.

References

Baccini, A., Laporte, N., Goetz, S. J., Sun, M., & Dong, H. (2008). A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters, 3*(4), 045011. https://doi.org/10.1088/1748-9326/3/4/045011

Brice, M., Wulder, M. A., White, J. C., & Hobart, G. (2013). Modeling stand height, volume, and biomass from very high spatial resolution satellite imagery and samples of airborne LiDAR. *Remote Sensing, 5*(5), 2308–2326. https://doi.org/10.3390/rs5052308

Dixon, R. K., Trexler, M. C., Wisniewski, J., Brown, S., Houghton, R. A., & Solomon, A. M. (1994). Carbon pools and flux of global forest ecosystems. *Science, 263*(5144), 185–190. https://doi.org/10.1126/science.263.5144.185

Dube, T., & Mutanga, O. (2015). Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in Umgeni catchment, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing, 101,* 36–46. https://doi.org/10.1016/j.isprsjprs.2014.11.001

Durkaya, B., Durkaya, A., Makineci, E., & Ülküdür, M. (2013). Estimation of above-ground biomass in different forest ecosystems of Turkey. *Forest Systems, 22*(2), 278–284. https://doi.org/10.5424/fs/2013222-02891

Günlü, A., Ercanlı, İ., 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. https://doi.org/10.15287/afr.2014.217

Güverçin, İ. (2024). *Saf kızılçam meşcerelerinde Sentinel-1A ve Landsat 8 OLI uydu görüntüsü kullanılarak topraküstü biyokütlenin tahmin edilmesi* [Yüksek lisans tezi, Çankırı Karatekin Üniversitesi].

Hirata, Y., Tabuchi, R., & Patanaponpaiboon, P. (2014). Estimation of aboveground biomass in mangrove forests using high-resolution satellite data. *Journal of Forest Research, 19,* 34–41. https://doi.org/10.1007/s10310-013-0402-5

Işık, S. (2024). *Saf sarıçam (Pinus sylvestris L.) meşcerelerinde Landsat 8 OLI uydu görüntüsü yardımıyla topraküstü karbon depolama miktarının belirlenmesi (Aladağ Orman İşletme Şefliği örneği)* [Yüksek lisans tezi, Kastamonu Üniversitesi].

Keleş, S., Günlü, A., & Ercanlı, İ. (2021). Estimating aboveground stand carbon by combining Sentinel-1 and Sentinel-2 satellite data: A case study from

Turkey. In *Forest resources resilience and conflicts* (pp. 117–126). Elsevier. https://doi.org/10.1016/B978-0-323-85368-8.00011-5

Lu, D. (2005). Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *International Journal of Remote Sensing, 26*(12), 2509–2525. https://doi.org/10.1080/01431160500142145

Orman Genel Müdürlüğü (OGM). (2024). *Orman İdaresi ve Planlama Dairesi Başkanlığı, Orman Amenajman Plan Verileri.* Ankara: OGM.

Saraçoğlu, Ö. (1988). *Karadeniz yöresi göknar meşcerelerinde artım ve büyüme.* Orman Genel Müdürlüğü Yayınları.

Sun, O., Uğurlu, S., & Özer, E. (1980). *Kızılçam (Pinus brutia Ten.) türüne ait biyolojik kütlenin saptanması.* OAE Yayınları Teknik Bülten Serisi No: 104.

Turgut, R. (2019). *Landsat 8 OLI uydu görüntüsü kullanılarak saf karaçam meşcerelerinde topraküstü biyokütlenin tahmin edilmesi* [Yüksek lisans tezi, Çankırı Karatekin Üniversitesi].

Zheng, D., Rademacher, J., Chen, J., Crow, T., Bresee, M., Le Moine, J., & Ryu, S. R. (2004). Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. *Remote Sensing of Environment, 93*(3), 402–411. https://doi.org/10.1016/j.rse.2004.08.008

Chapter 3

Türkiye's Ecological Independence: A Future in Harmony with Nature

Günay ÇAKIR[1], Fatih SİVRİKAYA[2], Nuri BOZALİ[3]

Introduction

Ecosystem services are the natural environment's circumstances, practices, and components that provide tangible and intangible benefits for meeting and sustaining human needs. A broad conceptual framework should be developed to handle the interactions between nature and humans (Daily, 1997; Diaz et al., 2015; Kandel et al., 2018). The environment is a source of several goods and services that raise the standard of living for people. Through the integration of goods and services, it provides individuals with social, ecological, and economic products (MEA, 2005; Daw et al., 2011; Reyers et al., 2013; Hicks et al., 2015). This human-centered viewpoint on nature promotes a new method for comprehending the influence of the environment on human welfare (Costanza et al., 1997; Daily and Matson, 2008; Scholes et al., 2013; Chaudhary et al., 2015). Every resource is available to and consumed by billions of individuals worldwide (Princen, 1999; French, 2000; Rees, 2010; Sivrikaya and Çakır, 2023). Because of the growing population, ecological processes have a more significant role in determining how people live.

Ecological independence has been identified as a notion in efforts to eradicate poverty in countries (Pereira et al., 2005; Grêt-Regamey et al., 2012; Sandhu and Sandhu, 2014; Suich et al., 2015; Kandel et al., 2018). The idea of ecological independence seeks to balance out the polar perspective. He believes there are several possible outcomes when making decisions about environmental monitoring. According to Michaelidou et al. (2002), ecological interdependence is defined as "*ecosystem preservation and*

[1] Karadeniz Technical University, Faculty of Forestry, Department of Wildlife Ecology and Management, Trabzon, Türkiye, Orcid:0000-0003-4951-4283, gcakir@ktu.edu.tr
[2] Kastamonu University, Faculty of Forestry, Department of Forest Engineering, Kastamonu, Türkiye, Orcid:0000-0003-0860-6747, fsivrikaya@kastamonu.edu.tr
[3] Karadeniz Technical University, Faculty of Forestry, Department of Forest Engineering, Trabzon, Türkiye, Orcid:0000-0001-5735-3649, nuribozali@ktu.edu.tr

community survival are two interconnected goals and equal focus should be given so that both can benefit." Thompson et al. (2011) state this unbreakable connection is *"the interdependence of living organisms with each other and in their physical environment with spatial and temporal differences"*. This concept lessens the existing emphasis on economic development and growth, especially in the context of environmental decision-making by people with the authority to influence critical social decisions (Harding et al., 2009).

Land encompasses the physical environment, comprising climate, topography, soils, hydrology, and vegetation, insofar as this affect land utilization. A land area is a geographically delineated section of the Earth's surface that includes the atmosphere, soil, underlying rocks, topography, water, flora and fauna, as well as the biosphere above and below it, all of which impact human utilization both presently and in the future. Consequently, land is a more expansive notion than only land and landmass, incorporating the concept of soil. Changes in ecosystems and landforms frequently account for variances in land mapping units within a site. Therefore, soil surveys are the fundamental tools utilized in defining and classifying land mapping units (Anonymous, 2008).

The psychological barrier caused by the disparity between costs and rewards over generations is the main hurdle for people, businesses, and governments adopting conservation practices. Costs must be paid in full during the so-called "critical decade," when benefits accrue to subsequent generations (Ehrlich and Ehrlich, 2013; Steffen and Hughes, 2013). Due to the lack of agreement on what causes environmental problems in the first place and the resulting disagreement over how best to spend money on environmental management, this barrier is detrimental to the environment (Adger and Jordan, 2009; Çakır et al., 2019).

Processes for achieving ecological independence are starting to take settlement areas, farms, pastures, the sea, terrestrial water surfaces, and forests into account. The development of information systems has accelerated across all industries because of technological advancement. Population information systems, business information systems, rangeland information systems, agricultural information systems, and information systems for managing water resources are only a few examples of information systems. It has been demonstrated in many other fields, such as how artificial intelligence has improved information systems.

It can be expanded to consider the various responsibilities that nations play in the global environment and how they frequently depend on one

another to keep it healthy, particularly in light of atmospheric issues like the ozone layer hole. According to Litfin (1999), Davis et al. (2009), and Thompson et al. (2011), community initiatives are a vital component of any conservation policy, whether it is being addressed at the local, regional, or global level. Environmental decision-makers of all sizes appear to confront a hurdle regarding including communities or individuals in conserving the natural resources we rely on (Newmark and Hough, 2000; Rist and Moen, 2013).

In this study, Türkiye's ecological independence was evaluated under the headings of development, population-education, environment, climate, agriculture, forest, and water resources.

Ecological Structure of Türkiye

Türkiye is situated between the 36°–42° north parallel and the 26°–45° east meridian in the northern hemisphere, at the geographic center of the continents of Asia, Europe, and Africa. A total of 814,578 km^2 is covered by the three separate climate types, including lakes and islands (Figure 1). The Black Sea, Mediterranean, and Aegean Seas all surround land masses with a temperate climate. In Türkiye, there are pure and mixed forests consisting of coniferous and deciduous trees. Because of Türkiye's geographic location, it has developed a rich variety of flora and fauna. The data used in this study were obtained from the Turkish Statistical Institute, General Directorate of Forestry, General Directorate of Meteorology, and other public agencies.

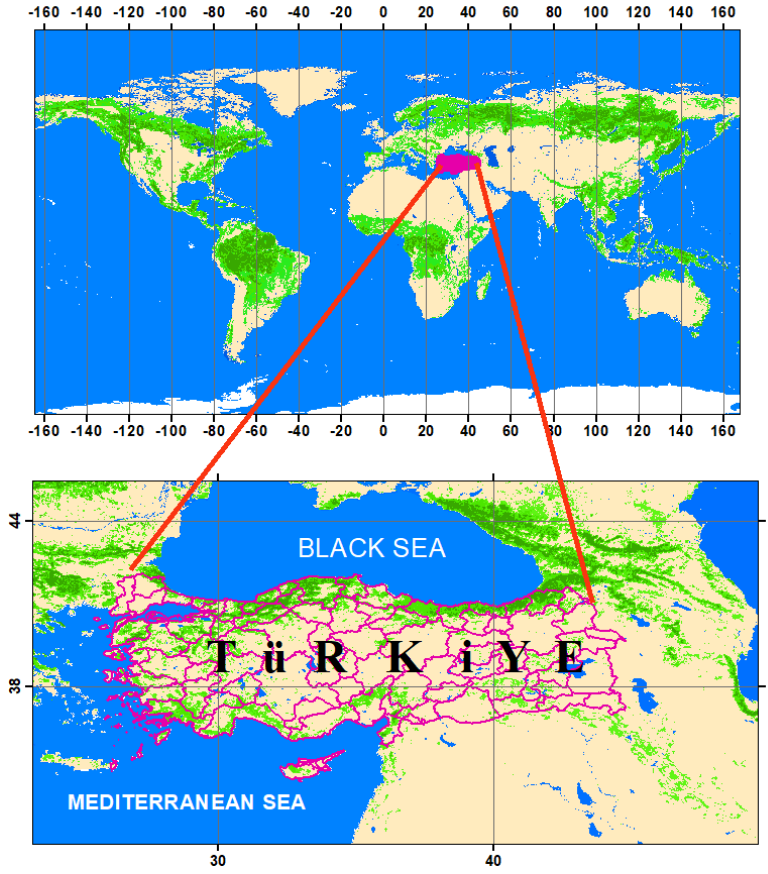

Figure 1. Study area

Table 1. General land layer changes in Türkiye over the years between 1990 to 2018 (Anonymous, 2024)

Years	Land Cover Types Area (Hectares)				
	Artificial surface	Agricultural areas	Forest and Semi natural areas	Wetlands	Water bodies
2018	1565407,01	34079354,82	40564303,45	413786,96	4013668,63
2012	1456764,95	34137732,58	40643583,57	412519,01	3985920,76
2006	1300589,00	33997989,52	40985565,14	417587,84	3938293,88
2000	1221002,63	33051010,40	42185865,01	283557,09	2847191,34
1990	962766,75	33483313,90	42114889,26	256308,21	2754732,50

According to Table 1, dams constructed in the past 30 years have contributed significantly to the rise in water surface area in Türkiye, which has resulted from changes in the country's general land use. The amount of forest and semi-natural areas has generally decreased, but the amount of artificial surfaces and agricultural areas has significantly increased (Anonymous, 2024).

Forty-six percent of Türkiye's land surface has a slope exceeding 40%. Furthermore, 62.5% of land above a 15% slope is affected by high slopes or topography. Due to high slopes or topography, agricultural areas face a significant erosion problem. General soil characteristics, climate, and socio-economic conditions also significantly impact land use types. Agricultural, forest, and pasture lands constitute the majority of Türkiye's land. When we look at land use in Türkiye, agriculture accounts for 38.7% of land displacement, forests 4.2%, and pastures 53.6%. Rainfall accounts for 14.3% of soil loss in Türkiye, while misuse of land accounts for 47.5% (Anonymous, 2019).

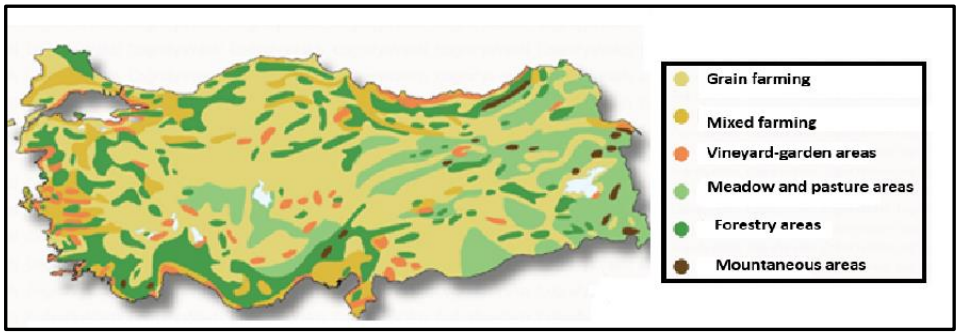

Figure 2. Türkiye's land use types (TUIK 2022)

Table 2. Detailed land use patterns of Türkiye between 1990 and 2018 (Anonymous, 2024)

	CORINE AREAS				
COVER TYPE	**Corine 2018 area (ha)**	**Corine 2012 area (ha)**	**Corine 2006 area (ha)**	**Corine 2000 area (ha)**	**Corine 1990 area (ha)**
Coniferous Forest (312) /Broad-Leaved Forest (311)/ Mixed Forest (313)	11525788,16	11621447,41	11638446,82	12200373,3	11713054,29
Land Principally Occupied by Agriculture, With Significant Areas of Natural Vegetation (243)	7137133,19	7146432,61	7264413,61	7607496,21	7756867,78
Sparsely Vegetated Areas (333)	9346868,30	9371380,33	9281650,83	10262845,63	10265069,77
Natural Grassland (321)	8877571,03	8907984,46	8931249,72	8996129,72	9153101,44
Transitional Woodland/Shrub (324)	7764016,10	7689048,55	7771647,09	7455919,45	7783000,22
Non-Irrigated Arable Areas (2111)/ Continuously Irrigated Areas (2121)/ Non-Irrigated Mixed Farm Areas (2421)/Non-Irrigated Fruit Fields (2221)/Olive Groves (223)	22850715,10	22867941,01	22796837,34	21844836,5	22021804,04
Pastures, Meadows and Other Permanent Grasslands Under Agricultural Use (231)	2009092,50	2041781,25	1975568,09	1491343,02	1632869,38
Irrigated Mixed Farm Areas (2422)/Irrigated Fruit Fields (2222)/Vineyards (221)/Rice Fields (213)/Greenhouse Areas in Continuously Irrigated Areas (2122)/Greenhouse Areas in Non-Irrigated Arable Areas (2112)/Annual Crops Associated with Permanent Crops (241)	2080474,79	2080880,23	1961170,48	2107334,65	2071772,69
Green Urban Areas (141)	11454,70	11493,56	10780,7	10455,99	10037,16
Coastal Lagoons (521)	17106,42	17167,02	15647,17	17818,75	
High Salt Content Bare Rocky (3322)	24534,65	24534,65	29479,26	136297,19	149467,98
Coastal Salt Marshes (421)	189411,52	189539,48	191113,73	26272,06	22335,75
Sclerophyllous Vegetation (323)	1061214,03	1065699,95	863071,88	922110,59	915271,15
Inland Marshes (411)	207571,35	208611,75	213215,34	250859	226539,29

Sea and Ocean (523)	2622084,25	2622878,51	2628768,21	1568091,07	1570290,94
Water Bodies (512)	1260803,95	1236411,72	1186547,72	1150669,8	1045649,94
Moors and Heathland (322)	164,31	164,31			
Bare Rocky (3321)	1867703,08	1868016,85	2360748,96	2070260,17	1991103,77
Water Courses (511)	113569,06	109358,55	107225,83	110337,58	121308,02
Agro-Forestry Areas (244)	1939,25	697,51			
Residential Areas Without Continuity (1122)	531103,41	520887,42	493824,87	656169,18	595716,59
Industrial or Commercial Units and Public Facilities (121)	222411,08	194433,53	165067,19	118142,73	59269,3
Beaches, Dunes and Sand Plains (331)	91898,28	92539,52	98309,11	132236,94	124060,67
Port Areas (123)	6385,42	6172,2	5574,65	3741,59	3372,94
Glaciers and Perpetual Snow (335)	626,99	626,99	589,92	589,92	650,71
Estuaries (522)	104,95	104,95	104,95	274,15	220,89
Continuous Urban Fabric (111)	136928,72	136271,69	123133,59	75939,55	60159,60
Mineral Extraction Sites (131)	142734,24	119295,47	95448,49	67214,30	42968,45
Construction Sites (133)	74939,53	57657,90	50301,76	33428,71	33187,75
Sports and Leisure Facilities (142)	43358,62	40661,70	35910,43	37221,81	21747,82
Road and Rail Networks and Associated Land (122)	42604,78	35538,94	28394,2	18183,24	2328,44
Airports (124)	31739,39	29458,15	27708,87	27255,69	21213,89
Salines (422)	16804,10	14367,77	13258,77	6426,02	7433,17
Dump Sites (132)	5580,31	4766,48	3528,94	2284,34	397,50
Burnt Areas (334)	3918,53	2140,55	10374,53	9102,14	17262,72
Rural Settlements Without Continuity (1121)	316166,80	300127,89	260915,31	170965,50	112367,29
Total	80636520,89	80636520,86	80640028,36	79588626,5	79551901,34

Development

The classification of a nation's development level also considers financial systems, energy, food security, and other areas. Natural resource usage is primarily concentrated in underdeveloped and developing countries (Daniel and Mittal, 2009). There could be a shortage of both land and water on a global scale as a result of the "global land grabbing" phenomenon. Food prices reflect the lack of natural resources, and markets experience a loss of confidence (IFPRI, 2009; UN, 2010). Developed nations purchase land to

meet food needs and create conservation regulations (Lipton, 2009). Carbon emissions rise as development levels do as well. The environmental effect values are increased by using fossil fuels to produce the required energy. This pollution is worsened by an excessive population living in cities. Low levels of carbon emissions are found in less developed nations. Nations do not fight against pollution because of their pollution levels (Okumus, 2020).

Population - Education

Türkiye's population grew from 66.4 million in 2002 to 86.0 million in 2022. The old population (65+ years) is increasing proportionately to the total population, whereas the population under the age of 14 is shrinking. The reason for this change in the share of age groups in the total population is that the total fertility rate has decreased below 2.1, which is the population renewal level, and human longevity has been extended due to advances in living standards in our country. The infant mortality rate is one of the demographic indicators. The infant mortality rate has reduced from 25 per thousand in 2002 to 9 per thousand in 2019 due to improvements in health care and living standards (TUIK, 2022).

Environmental education takes the form of educating people about the environment, helping them develop the necessary skills, and establishing lifelong habits. In Türkiye, environmental education happens on its own. As a result, the appropriate level of access could not be obtained. To increase society's understanding of the environment, efforts must be undertaken in several areas (Çolakoğlu, 2010).

In a study, more training with an environmental focus is targeted for delivery during particular days and weeks (Gülay and Ekici, 2010). Previous experiences have demonstrated that schooling removed from nature has a low success rate. Environmental education that occurs exclusively in a classroom does not suffice to increase ecological awareness (Özdemir, 2010). Environmental education should begin at home at a young age. At the moment of acquiring behavior, it is critical to provide education from childhood. These attitudes and behaviors serve as the foundation for desired behaviors in the future (Ateş and Karatepe, 2013).

Table 3. Number of people at education level between 2010-2024 (MEB, 2025)

Years	Primary / Secondary School	High School	University	Other	Total Population
2024	35329926	19016513	15100926	9950685	79398050
2020	36549582	15773910	12353732	10233350	79398050
2015	36417831	12990847	9149566	11939539	70497783
2010	29657224	11374336	5045702	20056199	66133461

Table 4. Number of students at different educational levels (MEB, 2025)

Academic Year	Preschool	Primary School	Secondary School	High School	University	Total population
2024-2025	1741314	5704483	5181914	5328812	6835115	79398050
2019-2020	1629720	5729945	5701564	5630652	7910133	79398050
2014-2015	1156661	5434150	5278107	5691071	6062886	70497783

As seen in Tables 3 and 4, Türkiye's contributions to education are quite substantial. However, after 2020, the number of students at different educational levels began to decline. This is a result of the country's declining fertility rate.

Although our country has a high level of environmental quality, a lack of education makes it difficult to appreciate its value. Nature conservation efforts are successful when the environment is kept clean. The cleanliness of the land, sea, and air can be attained not just on a national level but also on a global one through international cooperation. Even if you live in a clean place, the quality of your surroundings will decline as the world grows more polluted.

Environment

Local ecological alterations might have superregional repercussions when local production and consumption lead to on-site ecological degradation, jeopardizing the sustainability of distant areas (Kissinger et al., 2011). This environmental change can be caused by indirect and direct forces (MEA, 2005). Land use and land cover changes, technological adaptation, physical inputs, introducing and removing species, and local resource usage are all direct factors. Indirect factors affecting drivers include demography, economics, geopolitics, technology, and cultural/religious issues. Global and international processes drive some drivers, while local factors influence

others. A considerable amount of study has concentrated on production-related harm resulting from poverty, institutional failure, or the expansion of global markets. In an interconnected globe, resource extraction is predominantly influenced by local conditions, and local ecological alterations (e.g., species extinction, declining water levels, soil erosion, major pollution incidents, habitat loss, etc.) can adversely affect human well-being in other regions (Elisha and Felix, 2020).

Due to its soil characteristics, geographical location, topography, and temperature, Türkiye is vulnerable to desertification/land degradation, erosion, and drought. Furthermore, the effects of desertification/land deterioration, erosion, and drought are becoming more severe due to climate change and human activity. Land destruction results in a drop or loss of biological/ecological and economic productivity in productive areas such as agricultural fields and pastures, and in natural regions such as forests, wetlands, steppe, and marquis/heathland. This situation hurts the living conditions and livelihoods of the population, particularly women and youth who rely on natural resources in desertification/land degradation areas. According to one of the most prominent studies on desertification/land degradation in Türkiye, the Desertification Sensitivity Map, 22.5% of Türkiye has high desertification sensitivity, while 50.9% has medium desertification sensitivity. Furthermore, the main factors in desertification/land degradation in Türkiye, according to the national scale desertification criteria and indicators determined in the Türkiye Desertification Model, are water, climate, soil, topography and geomorphology, land cover, land use, socio-economics, and management (Anonymous, 2019).

Climate

Türkiye is situated between the temperate and subtropical zones. Surrounded by sea on three sides, the presence of mountains and the diversity of landforms have resulted in the emergence of various climate types (Kara et al., 2023). Türkiye's coastal locations have milder climate characteristics due to the impact of the water. The North Anatolian Mountains and the Taurus Mountains prevent seawater from accessing the interior. As a result, the heartland of Türkiye exhibits continental climate characteristics. According to worldwide climatic categorization criteria, Türkiye has the following climate types: continental climate, Mediterranean climate, Marmara (transitional) climate, and Black Sea climate (Atabay, 1997).

Table 5. The average temperature during 10-year periods, as the General Directorate of Meteorology reported. (MGM, 2021)

Years	Temperature Averages
Average Temperature in Türkiye Between 1970-1979	12.8°C
Average Temperature in Türkiye Between 1980-1989	12.7°C
Average Temperature in Türkiye Between 1990-1999	13.0°C
Average Temperature in Türkiye Between 2000-2010	13.6°C
Average Temperature in Türkiye Between 2011-2021:	14.1°C

Agriculture

Agricultural national income climbed from 37 billion TL in 2002 to 275 billion TL at the end of 2019. Agriculture's proportion of total employment is declining. Total agricultural and food product foreign trade increased from 6.8 billion dollars in 2002 to 31.4 billion dollars in 2019. Furthermore, overall exports of farming and food items exceed total imports. Total agricultural support payments grew from 2002 to 2020. Supports grew by 29% in 2020 compared to the previous year, totaling 21.97 billion TL. With the investments, 53000 hectares of land were made available for irrigation in 2019. Total livestock (cattle, buffalo, sheep, and goats) grew from 41.9 million in 2002 to 66.4 million in 2019. Total red meat output increased from 764 thousand tons in 2002 to 1.2 million tons in 2019, and total milk production increased from 8.4 million tons in 2002 to 23 million tons in 2019 (TUIK 2022).

Forest

Forest areas encompass over one-third of Türkiye's terrestrial area and hold significant ecological and economic value. Temporal variations in forest assets must be evaluated not solely by the expansion or reduction of forest area, but also by alterations in forest growing stock, annual growth potential, and species composition.

The first comprehensive assessment of Türkiye's forest assets occurred in 1973, revealing a forest area of 20.2 million hectares (GDF, 2020). Due to afforestation, erosion control, and regeneration initiatives implemented over subsequent decades, forest areas increased consistently; it constituted 26.7% of the nation's surface area in 1999 and rose to 29.8% by 2022 (GDF, 2022). In this timeframe, the overall forest area expanded by roughly 3.0–3.2 million hectares, attaining 23.2–23.3 million hectares. As of 2024, the forest area was 23.4 million hectares (GDF, 2025).

Data from the General Directorate of Forestry indicates that the nation's overall forest areas in 1973 were roughly 935 million m³, with an annual increment of about 28 million m³. As of 2024, the total volume amounted to 1,798,061,769 m³, with a yearly increment of 51 million m³ (GDF, 2025)

The tree species composition of Türkiye's forest predominantly consists of natural forests, showcasing a significant range of coniferous and deciduous trees. The predominant species include calabrian pine (*Pinus brutia*), black pine (*Pinus nigra*), Scots pine (*Pinus sylvestris*), beech (*Fagus orientalis*), oak (Quercus spp.), cedar (*Cedrus libani*), spruce (*Picea orientalis*), fir (Abies spp.), and juniper (Juniperus spp.). Oak species are projected to constitute around 29% of the nation's forest resources, red pine 23%, black pine 18%, and beech 8% (GDF, 2025).

Table 6. The spatial change of Türkiye's forest assets over the years (OGM, 2022)

Years	Total area	Productive		Degraded	
	Hectare	Hectare	%	Hectare	%
1973	20199296	8856457	44	11342839	56
1999	20763248	10027568	49	10735680	51
2005	21188747	10621221	50	10567526	50
2009	21389783	10972509	51	10417274	49
2010	21537091	11202837	52	10334254	48
2012	21678134	11558668	53	10119466	47
2015	22342935	12704148	57	9638787	43
2018	22621935	12983148	57	9638787	43
2019	22740297	13083510	58	9656787	42
2020	22933000	13264429	58	9668571	42
2021	23110000	13500000	58	9610000	42

Water

Earth's entire water volume is 1400 million km^3. Freshwater resources account for about 2.5% (35 million km3) of total water volume. Anonymous (2018) estimates that 68.7% (24 million km^3) of this amount is found as ice and snow in mountainous regions, Antarctica, and the Arctic, and 30.1% (8 million km^3) as groundwater. This sum is not divided evenly over the planet. Unbalanced water distribution causes;

-Rapid and uncontrolled population growth,

-Wasted water in various places based on the countries' progress,

-Technological advancements and industries are damaging water resources (lack of filtering and cleaning facilities),

- The negative consequences of changing climatic conditions on the existence of water resources can be counted.

There has been a water crisis due to inadequate water supplies in many parts of the world. The plan still includes numerous statements regarding upcoming water wars. The proportionate distribution of water usage shows that 69% is used for agriculture, 19% for industry, and 12% for households. Even less than 1% of the beverage and food industry, which includes bottled water, uses fresh water (Ma et al., 2023). According to Anonymous (2018), most of the 69% of water utilized for agriculture loses productivity due to unintentional consumption.

The hazards associated with Türkiye's climate and water supply are indicated in the official reports for 2020, and do not provide a favorable image of the rain pattern and water scarcity. These changes have been most noticeable since 2018, even if we did not notice them much in earlier years. The 11th Development Plan, *"Water Resources Management and Security Specialization Commission Report," is the result of the endeavor to serve as the foundation for the policies that the 2023 vision will form. This study discusses the measures and actions that can be implemented to mitigate the negative effects that may emerge in the next years. According to this assessment, "* in terms of our country's water resources, with 1500m^3 of usable water per person, it falls into the category of countries with water restrictions." As a result, the protection of existing water resources, the development of alternative water resources, the use or intensification of technologies compatible with the ecological environment, the development of sustainable water policies, and thus the creation of sustainable water policies, all have an essential place on the country's agenda as an input to sustainable development (Anonymous, 2018).

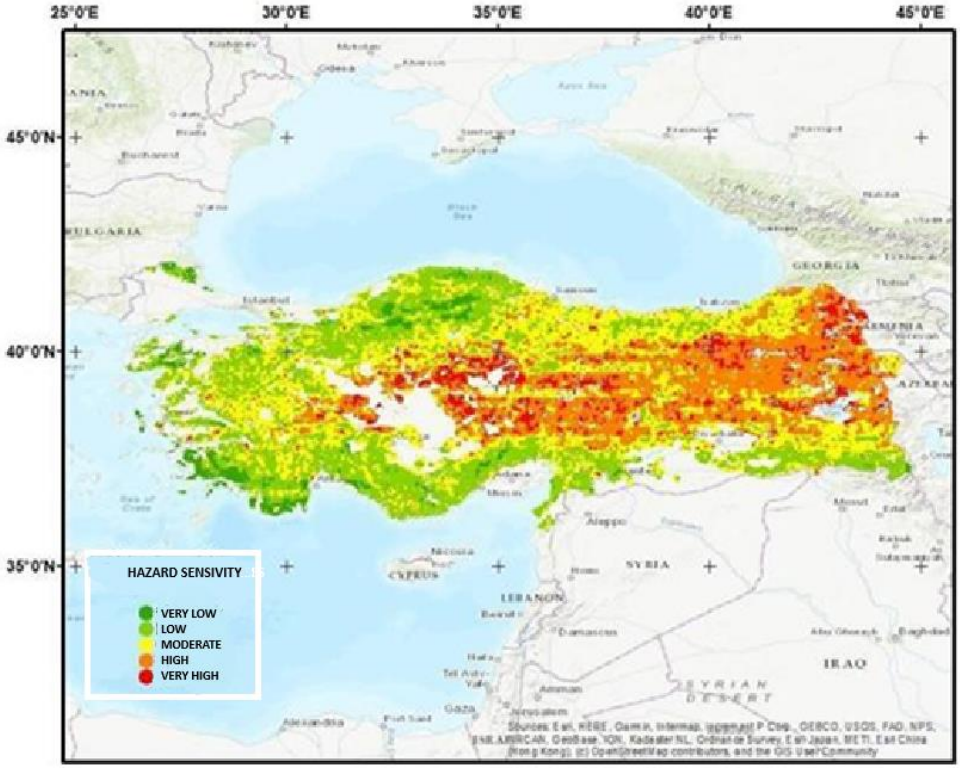

Figure 4. Water risk map turret (Anonymous, 2018)

Analysis of satellite images over a 25-year period showed no appreciable change in the amount of snowfall in the studied area. In the three provinces that comprise the Eastern Black Sea region, we are unable to pinpoint any significant effects of climate change on the water cycle. Nonetheless, it will be essential to stay up to date with advancements in the upcoming years. Given that forests are an essential part of the water cycle, the region's most notable feature is its abundant green tissue, which has delayed the rate of change (Çakir et al., 2025).

Conclusion

Because ecological viability and community survival are closely intertwined, successful conservation project design, monitoring, and assessment must focus equally on both (Harding et al., 2009). The literature on ecological independence suggests that we must balance environmental and social interdependence in this critical decade; if this concept is taken seriously immediately and effectively in all sectors, it will have far-reaching

consequences in all aspects of life in every corner of the world (Litfin, 1999; Çakir et al., 2008).

We design a new framework for understanding complex system failures and shaping more effective responses by articulating how countries and regions on a populated planet are interconnected through production and consumption, subsequent ecological change, and the unintended consequences of environmental policy responses. Due to rapid global transformation, the international community must urgently embrace a holistic Earth perspective, or at minimum, a superregional viewpoint, to align countries' shared interests in establishing new transnational institutions for multi-scalar development policy and decision-making. The essence of nature is clear: humanity must learn to coexist more fairly within the limits set by the ecosphere's capacity for renewal. Operating within these limitations necessitates a long-term perspective and a comprehensive, adaptive methodology. Regional and global development planning should account for biophysical constraints and authentic interregional ecological interdependence. Responding to the different interregional linkages detailed in this document should become a critical component of efforts to attain global sustainability.

Inter-sectoral collaboration helps ensure the healthy operation of ecological systems. In the global life cycle, all countries share the same level of responsibility. Türkiye has meticulously completed its tasks and continues to fulfill its obligations. Forest lands have grown as a result of planned forestry studies. It has contributed considerably to the carbon cycle, and environmental policies are progressively being implemented. It has been ensured that carbon footprints are lowered on solid grounds and that all options for the country's security have been utilized.

The global warming scenario has sparked international concerns about environmental security. However, it is becoming obvious that not all environmental security stakeholders have achieved a consensus. Despite differing approaches to climate and environmental security, governments do not completely agree on the problem. The traditional method implies that the global greening strategy is consistent. As a result, averting environmental disasters and coordinating national policies with this plan are critical. For example, it is difficult to determine settlements in places where sea level rise will be an issue. In reality, the message is clear all across the world: if climate change threatens national security, conserving the environment will eliminate that threat. In summary, all countries should support this process and contribute to a shared policy on this problem.

References

Adger, W. N., & Jordan, A. (2009). *Governing sustainability*. Edited by W. Neil Adger and Andrew Jordan, Cambridge, UK; New York: Cambridge University Press.

Anonymous, (2008). Toprak Ve Arazi Sınıflaması Standartları Teknik Talimatı ve İlgili Mevzuat. Tarım ve Köyişleri Bakanlığı Tarımsal Üretim ve Geliştirme Genel Müdürlüğü, Ankara, 192 s.

Anonymous, (2018). 11. Kalkınma Planı "Su Kaynakları Yönetimi Ve Güvenliği Özel İhtisas Komisyonu Raporu", T.C. Kalkınma Bakanlığı Yönetim Hizmetleri Genel Müdürlüğü Bilgi Ve Belge Yönetimi Dairesi Başkanlığı Kalkınma Bakanlığı Yayınları, Ankara.

Anonymous, (2019). Çölleşmeyle Mücadele Ulusal Stratejisi ve Eylem Planı 2019-2030, *Çölleşme ve Erozyonla Mücadele Genel Müdürlüğü Yayınları*, Ankara

Anonymous, (2024). Tarım ve Orman Bakanlığı Corine Projesi. https://corine.tarimorman.gov.tr/corineportal/index.html (01.05.2025)

Atabay, S., (1989). *Peyzaj Planlama, Şehir ve Bölge Planlama Bölümü Ders Kitabı*, YTÜ Matbaası, SP-372.18.89. İstanbul.

Ateş, M. & Karatepe, A. (2013). The Analysis of University Students' Perceptions towards "Environment" Concept with the Help of Metaphors. *International Journal of Social Science,* 6 (2), 1327–1348

Çakir, G., Bozali, N. & Sivrikaya, F. (2025). Snow cover detection using remote sensing techniques over different climate zones of Türkiye. Scientific Reports 15, 23993 https://doi.org/10.1038/s41598-025-07158-4

Çakir, G., Kaya, L.G., Yücedağ, C., Bayram, S.,2019. Determination of the Effects of Alucra Forest Planning Unit's Population Dynamics on Land Use Changes. Kastamonu Unıversıty Journal of Forestry Faculty, 19(1): 35-46. Doi:10.17475/kastorman.543415.

Çakir, G., Sivrikaya, F. & Keleş, S. (2008). Forest cover change and fragmentation using Landsat data in Maçka State Forest Enterprise in Turkey. *Environ Monit Assess* **137**, 51–66. https://doi.org/10.1007/s10661-007-9728-9

Chaudhary, S., A. McGregor, D. Houston, & N. Chettri. (2015). The evolution of ecosystem services: a time series and discourse-centered analysis, *Environmental Science & Policy*, 54:25–34.

Çolakoğlu, E. (2010). Haklar söyleminde Çevre Eğitiminin Yeri ve Türkiye'de Çevre Eğitiminin Anayasal Dayanakları. *Türkiye Barolar Birliği Dergisi,* 88, 151–171.

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., & Van Den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, *387*(6630), 253-260.

Daily, G. (1997). *Nature's services*. Island Press, Washington, D.C., USA

Daily, G.C., Matson, P. A. (2008). Ecosystem services: from theory to implementation. *Proceedings of the National Academy of Sciences USA* 105:9455–9456.

Daniel S. Mittal A. 2009. The Great Land Grab Rush for World's Farmland Threatens Food Security for the Poor. The Oakland Institute.

Davis, J. L., Green, J. D. & Reed, A. (2009). Interdependence with the environment: Commitment, interconnectedness, and environmental behavior. *Journal of Environmental Psychology*, 29, 173-180.

Daw, T., K. Brown, S. Rosendo, & R. Pomeroy. (2011). Applying the ecosystem services concept to poverty alleviation: the need to disaggregate human well-being. *Environmental Conservation* 38,370–379.

Díaz, S., Demissew, S., Joly, C., Lonsdale, W. M., & Larigauderie, A. (2015). A Rosetta Stone for nature's benefits to people. *PLoS Biology*, *13*(1), e1002040.

Ehrlich, P. R., & Ehrlich, A. H. (2013). Can a collapse of global civilization be avoided? *Proceedings of the Royal Society B: Biological Sciences*, *280*(1754), 20122845.

Elisha, O. D., & Felix, M. J. (2020). The loss of biodiversity and ecosystems: a threat to the functioning of our planet, economy and human society. International Journal of Economics, Environmental Development and Society, 1(1), 30-44.

French, H. (2000). *Vanishing Borders – Protecting the Planet in the Age of Globalization*. W.W. Norton & Company, New York.

GDF, (2022). Ormancılık istatistikleri, *Orman Genel Müdürlüğü Yayınları*, Ankara.

GDF, (2025). Available online: https://www.ogm.gov.tr/tr/ormanlarimiz/Turkiye-Orman-Varligi), (accessed on 23 July 2025).

Grêt-Regamey, A., Brunner, S. H., Kienast. F. (2012). Mountain ecosystem services: Who cares? *Mountain Research and Development,* 32(S1): 23–S34.

Gülay, H. & Ekici, G., (2010). MEB Okul Öncesi Eğitim Programının Çevre Eğitimi Açısından Analizi. *Türk Fen Eğitimi Dergisi*, 7(1), 74–84.

Harding, R., Hendriks, C.M. & Faruqi, M. (2009). Environmental decision-making: exploring complexity and context / Ronnie Harding, Carolyn M. Hendriks, Mehreen Faruqi, Annandale, N. S. W. Federation Press.

Hicks, C. C., Cinner, J. E., Stoeckl, N., McClanahan, T. R., (2015). Linking ecosystem services and human-values theory. *Conservation Biology*, 29: 1471–1480.

IFPRI, (2009). International Food Policy Research Institute, Climate Change Impact on Agriculture and Costs of Adaptations. Food Policy Report, Washington DC.28 pp.

Kandel, P., D. Tshering, K. Uddin, T. Lhamtshok, K. Aryal, S. Karki, B. Sharma, and N. Chettri. 2018. Understanding social–ecological interdependence using ecosystem services perspective in Bhutan, Eastern Himalayas. Ecosphere 9(2):e02121. 10.1002/ecs2.2121

Kara, Y., Akalin, I. A., Deniz, N. G., Dinç, U., Ünal, Z. F., & Toros, H. (2023). Utilizing ANN for improving the WRF wind forecasts in Türkiye. Earth Science Informatics, 16(3), 2167-2186.

Kissinger, M., Rees, W. E., & Timmer, V. (2011). Interregional sustainability: governance and policy in an ecologically interdependent world. Environmental science & policy, 14(8), 965-976.

Lipton, M., (2019). Land Reform in Developing Countries Property Rights and Property Wrongs. ISBN: 9780203876211. Taylor & Francis e-Library. 177 p.

Litfin, K.T. (1999). Constructing Environmental Security and Ecological Interdependence. *Global Governance*, 5: 359

Ma Y. Li Y.P. Huang G.H. Zhang Y.F. (2023). Sustainable management of water-agriculture-ecology nexus system under multiple uncertainties. Journal of Environmental Management. Volume 341, 118096.

MEA, (2005). Millennium Ecosystem Assessment – Ecosystems and Human Well-being: Synthesis. Island Press, Washington Available at: http://www.maweb.org/documents/ document.356.aspx.pdf (accessed April 2011).

MEB, (2025). The Ministry of Education, Academic Year Formal Education Statistics. Ankara. "https://istatistik.meb.gov.tr/OgrenciSayisi/EgitimKademelerineGore"

MGM, (2021). Meteoroloji Genel Müdürlüğü, *Meteoloroloji İstatistiki Bilgileri*, Ankara,

Michaelidou, M., Decker, D.J. & Lassoie, J.P. (2002). The Interdependence of Ecosystem and Community Viability: A Theoretical Framework to

Guide Research and Application. *Society & Natural Resources*, 15, 599-616.

Newmark, W. D. & Hough, J. L., (2000). Conserving Wildlife in Africa: Integrated Conservation and Development Projects and Beyond. *BioScience,* 50, 585-592.

Okumuş, İ. & Bozkurt, C. (2020). Ekonomik Büyümenin Çevreye Etkilerinin Farklı Gelişmişlik Düzeyindeki Ülkeler İçin İncelenmesi, Gaziantep University Journal of Social Sciences, 19(1), 238-255.

Özdemir, O., (2010). Doğa Deneyimine Dayalı Çevre Eğitiminin İlköğretim Öğrencilerinin Çevrelerine Yönelik Algı ve Davranışlarına Etkisi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi,* 27, 125–138.

Pereira, E., Queiroz, C., Pereira, H., and Vicente, L. (2005). Ecosystem services and human well-being: a participatory study in a mountain community in Portugal. *Ecology and Society* 10:14.

Princen, T. (1999). Consumption and environment: some conceptual issues. *Ecological Economics* 31, 347–363

Rees, W.E. (2010). Globalization and extended eco-footprints: neo-colonialism and (un)sustainability. In: Engel, J.R., Westra, L., Bosselmann, K. (Eds.), *Democracy, Ecological Integrity and International Law. Cambridge Scholars Publishing*, Newcastle (Chapter 24)

Reyers, B., R. Biggs, G. S. Cumming, T. Elmqvist, A. P. Hejnowicz, & S. Polasky (2013). Getting the measure of ecosystem services: a social-ecological approach. *Frontiers in Ecology and the Environment*, 11:268–273.

Rist, L. & Moen, J. (2013). Sustainability in forest management and a new role for resilience thinking. *Forest Ecology and Management*, 310, 416-427.

Sandhu, H., & S. Sandhu. (2014). Linking ecosystem services with the constituents of human well-being for poverty alleviation in eastern Himalayas, *Ecological Economics*, 107: 65–75.

Scholes, R.J., B. Reyers, R. Biggs, M.J. Spierenburg, and A. Duriappah. (2013). Multi scale and cross-scale assessments of social–ecological systems and their ecosystem services. *Current Opinion in Environmental Sustainability*, 5:16–25.

Sivrikaya, F., and Çakır, G. 2023. Assessing Spatio-Temporal Change and Dynamics of Forest Ecosystem Succession Using Patch Analysis. Kastamonu University Journal of Forestry Faculty. 23(3): 186-198. 10.17475/kastorman.1394879

Steffen, W. & Hughes, L. (2013). The Critical Decade 2013: Climate change science, risks and response. Commonwealth of Australia (Department of Industry, Innovation, Climate Change, Science, Research and Tertiary Education).

Suich, H., C. Howe, & G. Mace. (2015). Ecosystem services and poverty all aviation: a review of the empirical links. *Ecosystem Services,* 12:137–147.

Thompson, J. D., Mathevet, R., Delanoë, O., Gil-Fourrier, C., Bonnın, M. & Cheylan, M. (2011). Ecological Solidarity As A Conceptual Tool For Rethinking Ecological And Social İnterdependence İn Conservation Policy For Protected Areas And Their Surrounding Landscape. *Comptes Rendus Biologies*, 334: 412-419.

TUIK, (2022). Türkiye İstatistik verileri, TUIK Ankara.

UN 2010. World Economic and Social Survey 2010 Retooling Global Development. Department of Economic and Social Affairs. 201pp.

Chapter 4

Application of CNN Architectures to ALS Data for Forest Basal Area Prediction

Hasan AKSOY[1,2], Enes CENGIZ[3]

Introduction

Forest ecosystems play a fundamental role in natural balances and climate patterns, not only through the products and services they have provided for centuries, but also through their contribution to global carbon storage (Aksoy and Gunlu, 2025). These ecosystems are responsible for storing more than 40% of the global terrestrial carbon stock (Lahssini et al., 2022). Along with the services they provide, forests are home to a rich diversity of animals and plants (Alici and Dalkilic, 2022). Many researchers emphasize that the continuity and sustainability of forests, which encompass all these responsibilities, can only be ensured through planned management (Hansen et al., 2013; Shang et al., 2019). In accordance with all these matters, obtaining accurate, reliable, and timely parameters for stands, which form the building blocks of forest ecosystems, is of critical importance for both sustainable forest management and multi-level environmental monitoring and planning processes (White et al., 2016; Aksoy, 2024). The determination of stand parameters has been carried out based on traditional forest inventories using local measurements from the past to the present (Kangas et al., 2018; Hawryło et al., 2020). These inventories measure parameters such as forest volume (V), basal area (BA), number of trees (N), and mean diameter (d) locally in forests, and estimates are made (Kangas et al., 2018). The BA parameter mentioned above is an important indicator for assessing field quality (Rodríguez et al., 2010), estimating stand productivity (Fu et al., 2017), and optimizing stand structure (Ruiz-Benito et al., 204; Zeng et al., 2024). Additionally, the BA of a tree at breast height is the most important parameter used to define size classes and thus derive forest dynamics such as resource

[1] Sinop University, Vocational School of Ayancık, Department of Forestry, Program of Forestry and Forest Products, Sinop, Türkiye. Orchid: 0000-0003-1980-3834, haksoy@sinop.edu.tr
[2] Natural Resources Institute Finland, Yliopistokatu 6B, FI-80100 Joensuu, Finland. ext.hasan.aksoy@luke.fi
[3] Sinop University, Ayancık Vocational School, Program of Electronics and Automation, Sinop, Türkiye. Orchid: 0000-0003-1127-2194, ecengiz@sinop.edu.tr

availability and competition (Oliver and Olive, 1996; Cade, 1997; Husch et al., 2002; Fraser and Congalton, 2021). In addition, the volumes of individual trees, which are also necessary for obtaining the total stand volume, are traditionally calculated by multiplying the cross-sectional area at breast height by the tree height and a form factor. Therefore, it is evident that the accurate determination of BA is a very important factor in obtaining accurate growing stock volume estimates (Husch et al., 2002; Witzmann et al., 2022). BA parameter measurements and calculations in forest inventory studies are generally based on the diameter at breast height (DBH), which is measured using calipers or a diameter tape, assuming a circular cross-section (Tischendorf, 1943; Witzmann et al., 2022). However, these traditional methods require a high amount of labor and time and can generally only be applied in limited areas (Dube et al., 2017). Additionally, the rapid change in forest stand parameters due to mechanisms such as anthropogenic climate change, land fragmentation, changes in land cover, and overuse has created a need for faster, more reliable, and large-scale data collection methods by integrating modern technologies with traditional forest inventory methods (Fraser and Congalton, 2021). In this context, remote sensing techniques have proven to be a valuable data source for forest management and inventory due to their ability to replicate a synoptic view of the Earth's surface and their potential for cost-effective utilization (McRoberts et al., 2010; Giri et al., 2011; Shang et al., 2019). The availability of spectral satellite imagery, particularly from Landsat and Sentinel, has encouraged researchers to focus on estimating forest parameters using spectral features obtained from the satellite imagery (Aksoy, 2024). Although there are many studies aimed at estimating forest parameters using a variety of advanced remote sensing techniques and different satellite images by researchers, optical remote sensing continues to lack the natural ability to measure the vertical structure of the forest and cannot be fully explained due to certain limitations (Mora et al., 2013; Shang et al., 2019).

Recent advances in airborne laser scanning (ALS) technology have significantly changed the role of remote sensing in estimating forest parameters, largely overcoming the limitations of optical remote sensing techniques. Numerous studies have demonstrated that both the identification of individual trees and the use of canopy height distribution characteristics at the plot level can produce highly accurate forest parameters (Maltamo et al., 2007; Kangas et al., 2019; Toivonen et al., 2021). ALS based approaches are LiDAR-based systems that provide 3D point clouds describing the structure of the vegetation cover from the canopy top to the ground. Therefore, the 3D spatial distribution of points resulting from laser-environment interaction can be directly related to the spatial distribution of vegetation cover (Lahssini et al., 2022). The combination of plot

or single tree based ground based forest parameter measurements and metrics derived from ALS point clouds with non-parametric modeling techniques enables the development of empirical models to estimate structural and biophysical forest variables (Bouvier et al., 2015; Véga et al., 2016).

Various statistical and machine learning based methods are used to estimate forest stand parameters in forest ecosystems. Traditionally, one of the most frequently used methods for this purpose has been multiple linear regression (MLR) (Kelek et al., 2021; Yıldız et al., 2023). However, in recent years, with the increase in data volume and diversity, more flexible methods capable of providing higher accuracy have begun to be preferred. In this context, machine learning algorithms such as k-nearest neighbors (k-NN), classification and regression trees (CART), artificial neural networks (ANN), random forest (RF), decision trees (DT), and support vector machines (SVM) offer important alternatives for estimating stand parameters (Xu et al., 2018; Zhao et al., 2019; Aksoy, 2024). In addition, with the development of deep learning approaches, convolutional neural networks (CNNs) have emerged as a powerful method for extracting complex patterns from high-dimensional data (Kelek et al., 2025; Özer et al., 2022). CNN algorithms enable the automatic processing of features obtained from spatial data, such as LiDAR and satellite images, and the more detailed modeling of forest structure. Therefore, the high representational power offered by CNNs has made them a more powerful alternative to classical machine learning methods for estimating forest parameters (Balazs et al., 2025).

Research on BA estimation has made significant progress using different remote sensing data and modeling techniques. Maltamo et al. (2007) modeled BA and diameter distribution using ALS data and achieved high accuracy (R^2=0.76) with a calibration-supported approach. Pippuri et al. (2013) combined ALS data with existing stand records in urban forests to estimate BA by species and obtained strong results, particularly in species differentiation (R^2=0.81). Shang et al. (2019) integrated ALS data with multi-seasonal satellite imagery to improve BA estimation accuracy and demonstrated the advantages of using data types together (R^2=0.83). He et al. (2021) established additive systems in natural oak forests, using breast height diameter and canopy stratification in basal area estimates, thereby increasing estimation consistency (R^2=0.79). Jevšenak and Skudnik (2021) estimated BA increase using the RF algorithm with national forest inventory data and reported that RF provided high accuracy (R^2=0.85). Fraser and Congalton (2021) contributed to forest inventory using UAS data and demonstrated the advantages of UAS imagery in BA estimation in small-scale areas (R^2=0.80). Witzmann et al. (2022) achieved high agreement in BA estimation by comparing TLS point clouds with diameter measurements

(R^2=0.88). Lahssini et al. (2022) integrated Sentinel-2 data with LiDAR metrics through ANN and demonstrated the strong estimation capabilities of ANN models in complex forest structures (R^2=0.87). Zeng et al. (2024) developed additive BA models in broadleaf mixed forests and provided more consistent estimates (R^2=0.82). Finally, Aksoy (2024) compared different machine learning algorithms using Landsat 8 OLI and Sentinel-2 images and showed that the best performance in BA estimation was achieved with the RF model (R^2=0.89).

As seen in the studies summarized above, different data sources and various modeling techniques have yielded successful results in BA estimation. However, it is observed that deep learning-based approaches, especially CNN, have been addressed in a very limited number of studies in the estimation of forest stand parameters. The CNN architecture has a strong potential to adapt to the high-dimensional and detailed structure of ALS data, thanks to its superiority in capturing spatial patterns and extracting meaningful features from complex data structures. In this context, the main objective of the current study is to estimate the basal area using ALS data through the CNN architecture and to evaluate the accuracy performance of this approach. Thus, it aims to contribute to the limited CNN based modeling studies in the literature and to present a new methodological perspective in BA estimation.

Material and methods

Study area

This research was conducted in boreal forest ecosystems located in the Mikkeli region of Finland. The study area is in eastern Finland between approximately 61°41′–62°11′ north latitude and 26°25′–27°27′ east longitude. The area, situated in the southern boreal biogeographic zone, exhibits similar characteristics in terms of climatic conditions, topography, and land use. The region has subarctic climate characteristics, classified as Dfc; annual average temperatures range from -6 °C to 17 °C, while extreme values are measured between -27 °C and +31 °C. Annual precipitation is approximately 740 mm, with an average of 147 days of rainfall. Mikkeli's topography is characterized by elevations ranging from 70 to 223 m and a gently undulating terrain with numerous lakes. The average elevation in Mikkeli is approximately 110 m, and the elevation differences in the environment generally do not exceed 100 m (Anonymous, 2023). The dominant species forming the forest cover in the study area consist of both coniferous and broad-leaved trees. Among the coniferous species, Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*) are particularly widespread, while among the broad-leaved trees, silver birch (*Betula pendula*) stand out as key components of the ecosystem. These species are the

main tree types that shape the ecological structure of the region and play a decisive role in forest dynamics (Balazs et al., 2025).

Ground measurements and ALS data

The field data used in this study were obtained by the Finnish Forest Centre (SMK) in 2020. In the sample areas, all tree individuals with a diameter at breast height (DBH) of 8 cm and above were recorded. The spatial information of the trees was determined using digital diameter gauges capable of interacting with pseudoliths. In order to create regular and comparable data sets, a total of 750 circular sample plots, each with a radius of 9 m (254.34 m²), were established. BA, used as the response variable in the analyses, was calculated for each sample plot and prepared for inclusion in the modeling process. The ALS data used in the study were collected in the same year as the ground measurements using a RIEGL VQ-1560i laser scanner at a 20% side overlap rate and a pulse density of 7.2 pulses/m², at a height of 1,525 m above ground level.

Methods

The general methodology of the study consists of three main stages. The first stage involves obtaining BA, which will be used as the response variable, through local measurements (inventory). The second stage involves processing the ALS data obtained to cover the study area and preparing the datasets by calculating the characteristics from the ALS data for each sample plot, calculating the independent variables to be used as inputs in the model. Finally, the third stage involves estimating BA using the CNN modeling technique with the prepared datasets and evaluating prediction accuracy using performance criteria. The general workflow diagram for the study is shown in Figure 1.

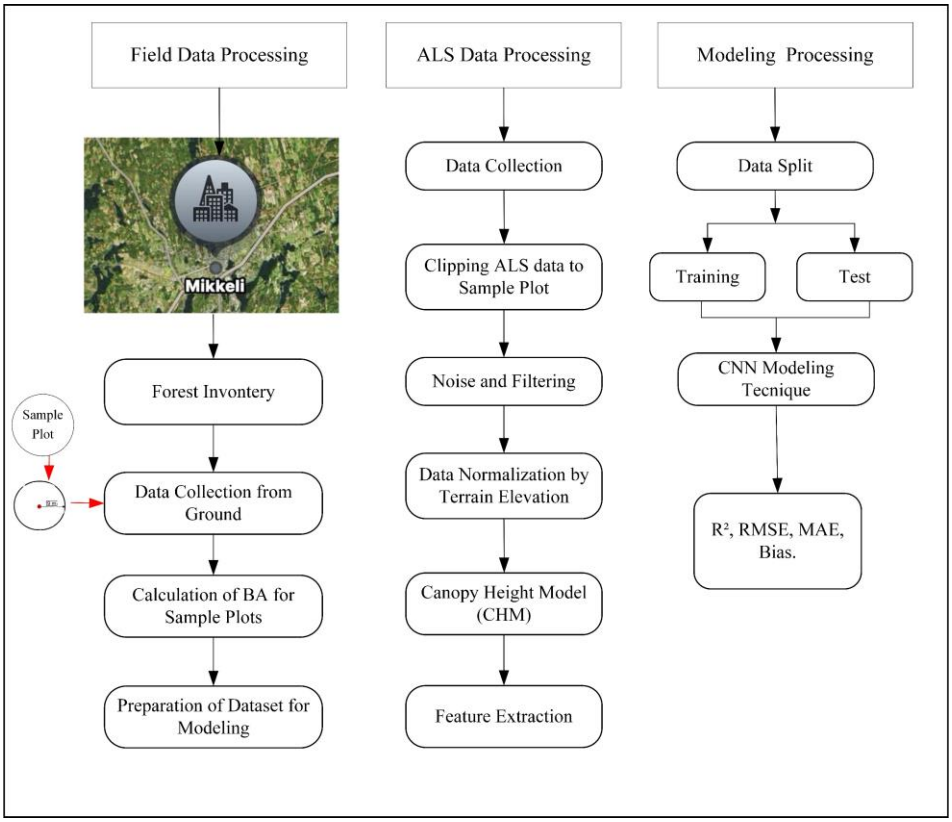

Figure 1. The generally flowchart of study

Airborne laser scanning data processing and features calculating

In this study, the R programming language was used to calculate all characteristics of the LAS data and sample plots. In the first stage, ALS data were clipped for each sample plot to obtain structural and spectral parameters related to forest areas. To determine ground elevation and separate vegetation height from topographic slope, the Digital Terrain Model (DTM) produced by the National Land Survey Agency (NLS) was used, and ALS data were normalized for each sample area. This aimed to determine the actual height values of trees in the plots. Furthermore, points that created noise in the dataset and fell outside the defined height limits were filtered out and excluded from the analysis. To create the Canopy Height Model (CHM), the ALS data for all sample plots were converted to raster format at a resolution of 0.4 m, and canopy height values were calculated using the point-to-raster (p2r) method. Then, a segmentation technique was applied to the CHM to identify the trees in each sample plot. The results obtained were compared with the tree locations recorded in the field, and an

accuracy assessment was performed. This process made it possible to accurately identify trees of different sizes.

Following all these steps, ALS based features were calculated for each sample plot. Basic distribution measures such as maximum, mean, standard deviation, skewness, kurtosis, and entropy were extracted within the scope of height statistics. In addition, different percentile intervals and quartiles (e.g., zq5, zq10, …, zq45) were calculated for the elevation values, and the elevation profile was analyzed in detail. As a result, ALS features to be used as independent variables were systematically derived for each observation (Figure 2).

CNN modeling technique

CNN1D Model: The CNN1D model consists solely of convolutional and fully connected layers, representing a pure convolutional approach. The initial Conv1D layer extracts 32 features from the input, followed by a Rectified Linear Unit (ReLU) activation and a Dropout operation (with a dropout rate of 0.3) to enhance generalization. The data are then flattened using the view function and passed through two fully connected (Linear) layers. The first fully connected layer contains 64 neurons, while the second produces a single scalar output. This structure is well-suited for time-series or sequential regression/classification tasks where long-term dependencies are not critical. Compared to recurrent models, CNN1D offers computational efficiency and effectively captures local temporal patterns.

CNN1D_LSTM Model: This model integrates a one-dimensional convolutional layer (Conv1D) with a Long Short-Term Memory (LSTM) layer, forming a hybrid architecture. Initially, the Conv1D layer extracts 32 feature maps from the input, with a kernel size of 3 and padding of 1 to preserve temporal dimensions. The ReLU activation function is applied to introduce nonlinearity. The extracted features are subsequently permuted and fed into an LSTM layer, which captures temporal dependencies within the sequence. The LSTM employs 32 hidden units and is configured with $batch\_first = True$ to process batches efficiently. The output from the LSTM is regularized using a Dropout layer (with a probability of 0.3) to prevent overfitting. Finally, the model employs a Linear layer with a single output unit for prediction. This architecture is particularly effective for tasks such as time-series forecasting or sequential data classification, where the CNN component captures local temporal patterns while the LSTM component models long-term dependencies.

CNN2D Model: The CNN2D model extends the convolutional approach to two-dimensional representations, enabling the network to capture spatial correlations within the input data. A Conv2D layer with 32 filters and a kernel

size of 3 is applied after reshaping the input into a two-dimensional structure. The side length of this reshaped input is determined as the ceiling of the square root of the input dimension. When the data length does not fit perfectly into a square shape, zero-padding is applied to complete the missing elements. Following the convolutional operation, the ReLU activation and Dropout (0.3) layers are used for nonlinearity and regularization, respectively. The output is then flattened and passed through two fully connected layers, resulting in a single scalar prediction. This architecture is particularly suitable for multidimensional sensor data or image-like time-series representations, where spatial structures and inter-feature dependencies are significant.

Modelling and model performance criteria

In this study, BA stand parameter was modeled using CNN modeling techniques with three different datasets. The accuracy performance of the modeling techniques was measured using the determination coefficient (R^2, Eq. 1), root mean square error (RMSE, Eq. 2), mean absolute error (MAE, Eq. 3), and bias (Bias, Eq. 4) criteria.

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

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

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (3)$$

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

Where n is the number of observations, k is the number of coefficients, y_i are measured the stand parameters, $\hat{y}_i$ are predicted stand metrics, $\bar{y}_i$ is a mean of measured stand metrics values, p is the number of parameters.

Results and Discussion

This section presents the performance results of the CNN1D, CNN1D-LSTM, and CNN2D models considered for BA estimation. The MAE, RMSE, R^2, bias, and last loss results of the models run for 100 epochs are given in Table 1.

Table 1. Performance results of the compared models

Model	R^2	RMSE	MAE	Bias	Last Loss
CNN1D	0.842	0.376	0.279	0.065	0.081
CNN1D-LSTM	0.673	0.541	0.427	0.115	0.368
CNN2D	0.860	0.354	0.265	0.058	0.047

When the table is examined collectively, the three architectures exhibit distinct behaviors regarding error magnitude, explained variance, and systematic bias. The coefficient of determination (R^2) values indicate the explanatory power of each model. CNN2D ($R^2 = 0.860$) achieved the highest proportion of explained variance, closely followed by CNN1D ($R^2 = 0.842$). In contrast, CNN1D-LSTM ($R^2 = 0.673$) demonstrated a notable decline in performance, implying that it was less effective in capturing the variance of the target variable. This difference suggests that CNN2D, by incorporating two-dimensional representations, was able to learn spatial and local correlations more effectively, whereas CNN1D despite its simpler design still captured sufficient information for strong performance. Adding an LSTM component may have increased model complexity without providing additional benefits, possibly due to insufficient temporal dependencies or overfitting tendencies in the available dataset. The error metrics (MAE and RMSE) provide additional insights into the distribution of residuals. For CNN2D, the close proximity of $\text{MAE} = 0.265$ and $\text{RMSE} = 0.354$ indicates relatively homogeneous errors and limited influence of extreme outliers. CNN1D shows a similar pattern ($\text{MAE} = 0.279$, $\text{RMSE} = 0.376$), though the slightly higher RMSE implies occasional larger deviations. In contrast, CNN1D-LSTM exhibits substantially higher errors ($\text{MAE} = 0.427, \text{RMSE} = 0.541$), and the larger gap between the two suggests the presence of more pronounced outliers. This pattern indicates that the LSTM component may have introduced instability, leading to heavier-tailed error distributions. The Bias values reveal the models' systematic tendencies toward over- or underestimation. All three models demonstrate a positive bias (CNN2D: 0.058, CNN1D: 0.065, CNN1D-LSTM: 0.115), implying a general tendency to overestimate the target variable. The smallest bias in CNN2D and the moderate value in CNN1D indicate relatively well-calibrated predictions, while CNN1D-LSTM's higher bias suggests miscalibration, consistent with its lower accuracy metrics. In practice, such systematic bias could be mitigated through post-hoc calibration or bias correction techniques. The Last Loss metric reflects the final state of the training process. Lower loss values for CNN2D (0.047) and CNN1D (0.081) indicate stable convergence and good generalization, while the notably higher loss in CNN1D-LSTM (0.368) points to unstable learning dynamics, possibly due to suboptimal hyperparameter tuning (e.g., learning rate, dropout, or hidden size) or inadequate

regularization. However, the last training loss alone cannot fully represent generalization capacity; it should ideally be interpreted alongside validation loss or cross-validation results.

In summary, the analysis reveals several important insights: (i) CNN2D effectively captures spatial and local dependencies, offering strong performance in both error and variance-based metrics, (ii) CNN1D achieves competitive results with lower architectural complexity and computational cost, and (iii) integrating an LSTM component without sufficient sequential structure or proper optimization may lead to performance degradation, as seen in CNN1D-LSTM. Figure 2 shows the scatter plots of the CNN1D, CNN1D-LSTM, and CNN2D models.

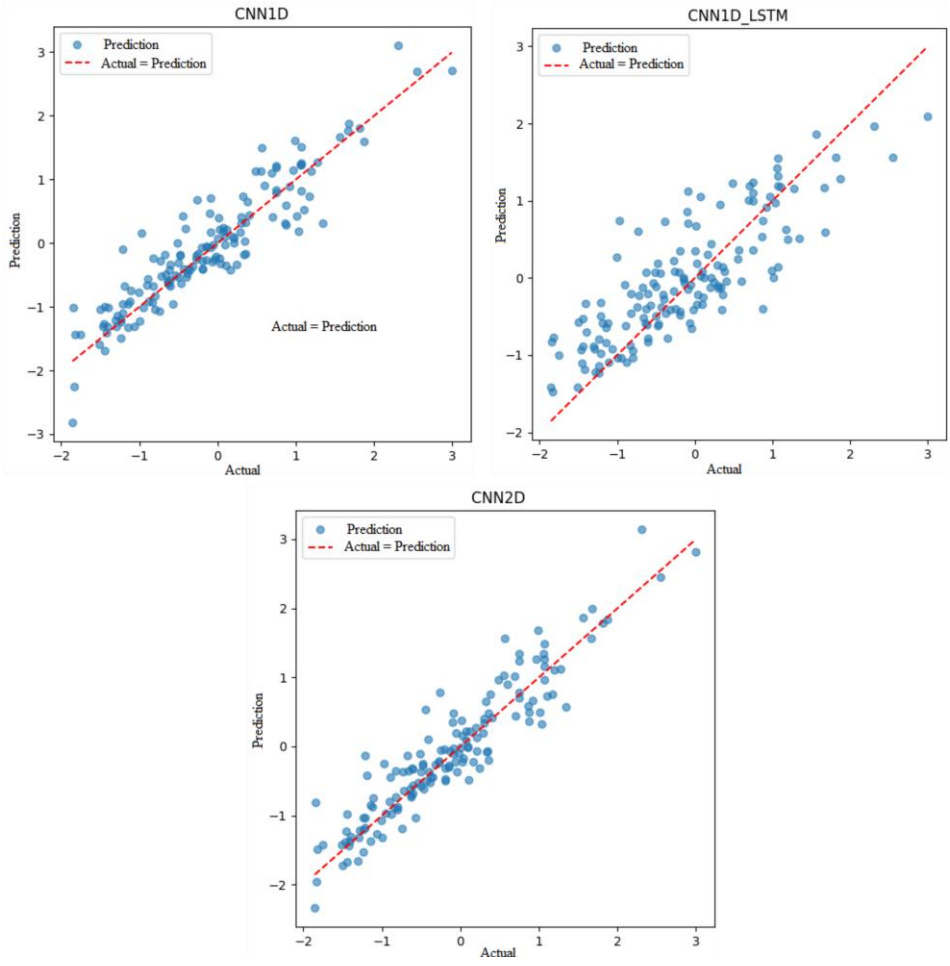

Figure 2. The scatter plots of the CNN1D, CNN1D-LSTM, and CNN2D models

The Figure 2, illustrating the predictive performance of the three models (CNN1D, CNN1D-LSTM, and CNN2D) reveal the relationship between the actual and predicted values of the BA variable. The red dashed line in each graph represents the ideal condition (Actual = Prediction), while the distribution of blue points around this line provides visual insights into each model's accuracy, calibration, and systematic errors. Overall, all three models display a positive linear relationship; however, the density and spread of the points around the line clearly reflect differences in model performance. In the CNN1D model, the predictions are generally positioned very close to the 1:1 line, successfully capturing the overall trend of the data. Most of the points are clustered near the red dashed line, with only minor deviations observed at the extreme values. This pattern indicates that the model exhibits high predictive accuracy and stability, though a slight margin of error is present for extreme observations. In contrast, the CNN1D-LSTM model shows a more dispersed distribution. Although the overall linear trend is maintained, the points are scattered more widely around the reference line. This dispersion indicates lower predictive accuracy and higher levels of both bias and random error. The noticeable deviations in the mid-value range suggest that the model tends to systematically overestimate or underestimate certain observations, implying calibration issues in specific regions of the prediction range. For the CNN2D model, the relationship between predicted and actual values appears the most consistent. The points are tightly clustered along the 1:1 line, with a balanced distribution across both low and high values. Youssef et al. (2022) similarly demonstrated that CNN2D performed 3.5% better than CNN1D on the dataset they used. This indicates that the two-dimensional convolutional structure effectively captures spatial patterns within the data, resulting in highly accurate predictions with minimal systematic bias. In general, the graphical findings align closely with the quantitative performance metrics (R^2, RMSE, MAE, and Bias).

Figure 3 shows the training loss curves of the CNN1D, CNN1D-LSTM, and CNN2D models.

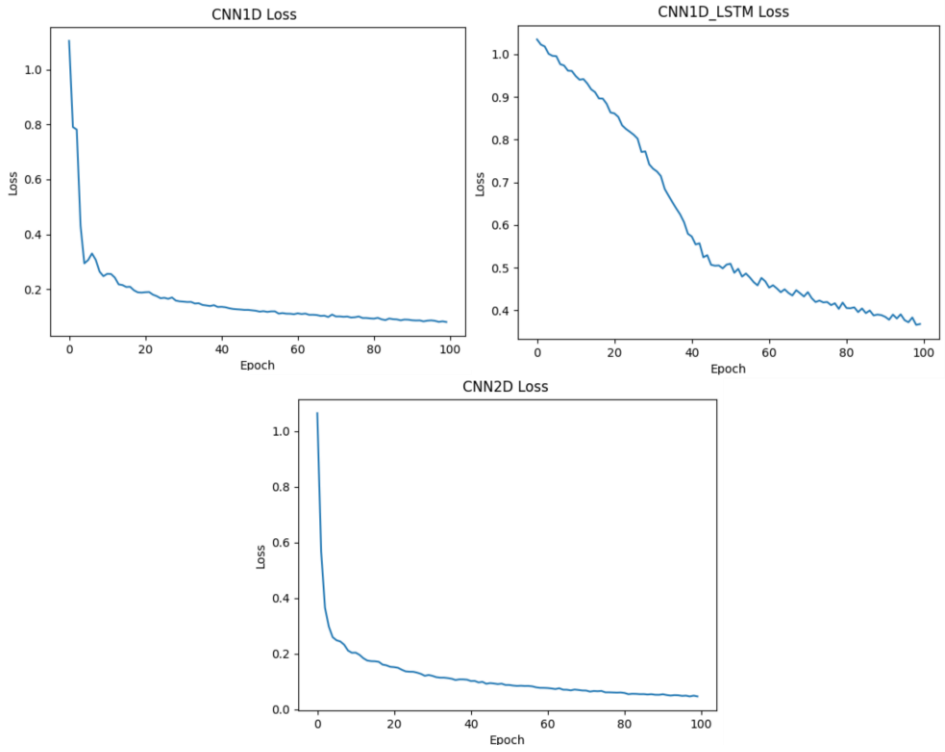

Figure 3. Training loss curves of the CNN1D, CNN1D-LSTM, and CNN2D models

Examining the loss graphs of the training processes of the three models (CNN1D, CNN1D-LSTM, and CNN2D) reveals valuable insights into their learning speeds, convergence behaviors and overall stability levels. Although the loss value in the CNN1D model was initially high, it declined rapidly within the first few epochs before slowing down and stabilizing at a low level. This suggests that the model achieved a balanced learning process quickly and was able to generalize in a stable manner. A similar but more effective learning curve was observed in the CNN2D model, which reduced the loss quickly in the early epochs and reached the lowest error level by the 100th epoch. This demonstrates that the CNN2D architecture successfully captures spatial patterns in the data and is highly efficient in parameter optimization. By contrast, the loss value in the CNN1D-LSTM model decreased more slowly, and the model ended with a relatively high loss value only after a long training process. This suggests that, although adding the LSTM layer provides flexibility in terms of time dependency, it negatively affects learning speed and efficiency. Overall, all three models successfully converged during the learning

process. However, the CNN2D model demonstrated the highest training performance, achieving the lowest final loss value and the most stable learning curve.

The overall results of the study (Bias=0.058) indicate that the model does not show a systematic tendency to overestimate or underestimate. This is thought to be due to the accuracy of the parcel ALS matching performed during the data preprocessing process. Janssen et al. (2023) similarly noted that most errors stem from coregistration errors between inventory parcels and remote sensing data. The limited extent of this error in our results indicates that the spatial alignment of the data is high, along with the structural characteristics of the model. Furthermore, it is thought that CNN2D gains tolerance to small spatial shifts by considering the relationships between neighboring pixels, thereby reducing the margin of error. The CNN1D-LSTM model's lower performance compared to other models (R^2=0.673) is thought to stem from the nature of the BA variable, which is spatial rather than temporal. Although the LSTM structure has advantages in learning temporal dependencies, the model's capacity could not be fully utilized in this case due to the absence of time series data. Furthermore, the high number of parameters in the model relative to the available data size may have led to overfitting. Therefore, it is thought that LSTM does not provide any additional benefit in such spatial data sets and may even reduce performance. The use of only ALS data in our study is noteworthy when compared to studies in the literature that utilize multi-source data integration. Lahssini et al. (2022) and Brown et al. (2022) have demonstrated that combining LiDAR data with optical data such as Sentinel-2 provides significant improvements in BA prediction. These studies emphasize that spectral information is related to characteristics such as stem density, leaf area, and chlorophyll content, especially in areas with high species diversity and structural complexity. Since our model uses only ALS data, the lack of this spectral information constitutes a limitation. Nevertheless, the strong results obtained by the CNN2D model suggest that learning spatial representation can partially compensate for this deficiency.

When the results obtained are evaluated in comparison to studies using classical machine learning based methods, Hosseini et al. (2025) compared various machine learning algorithms in *Fagus orientalis* stands using national inventory data and noted that tree-based methods (RF, GBM) generally offer strong baseline performance. However, the R^2 value achieved by our CNN2D model is close to or higher than the values obtained with such methods. This suggests that spatial patterns can be effectively captured by deep learning models and that the CNN structure is suitable for a spatially dependent variable such as BA. However, as Hosseini et al. (2025) also emphasize, the generalizability of such models needs to be tested across different regions and species. Therefore, it is important to test our model on different forest types. As a result, it is thought that the superior performance of the CNN2D

model is related to BA being a spatially organized variable. The model's low error and bias values can be explained by both the high spatial resolution of the ALS data and the ability of 2D convolutional structures to extract meaningful patterns from this data. However, the model's generalizability can be improved by integrating optical data and conducting larger geographical tests. Overall, our findings support the critical role of accurately modeling structural indicators representing BA distribution balance in understanding forest dynamics, as emphasized in Adnan et al. (2022)'s BALM approach. In conclusion, CNN-based spatial models offer a powerful alternative for BA estimation from ALS data, but they can be further improved with multi source data fusion and regional validity tests, as suggested in the literature.

Conclusions

In this study, the performance of deep learning based models developed to predict BA using ALS data was examined. The results showed that the CNN2D model offered the highest accuracy (R^2=0.860) and low error values (RMSE=0.354, MAE=0.265) compared to other models. Although the study was limited to ALS data, it demonstrated that such data can be processed by deep learning models to form meaningful patterns. However, it is thought that multi source data fusion with optical data could further improve the model's accuracy in the future. The literature emphasizes that the combination of multispectral data and LiDAR provides significant contributions to BA prediction. These findings should be validated with larger datasets in future studies. Future studies should primarily include the integration of optical data (such as Sentinel 2) with ALS data. This fusion could provide additional benefits for high accuracy BA estimates. Furthermore, testing the model's generalizability across different forest types and geographic regions will increase its robustness and scope. To further improve the model's performance, it is recommended to increase the accuracy of co-registration processes and test on larger datasets. Furthermore, applying class-balanced training methods and quantile loss functions to better model outliers is expected to reduce errors, particularly in high BA areas.

Acknowledgments

This study was also conducted during the author Hasan AKSOY's tenure as a visiting scientist at the Finnish Natural Resources Institute (LUKE) under the TÜBİTAK 2219 Postdoctoral Research Fellowship program (Grant No: 1059B192302166). Author Hasan AKSOY expresses his gratitude for the support provided by TÜBİTAK. In addition, the authors sincerely thank the Natural Resources Institute Finland (LUKE) for its support in providing the data.

References

Adnan, S., Valbuena, R., Kauranne, T., Gopalakrishnan, R., & Maltamo, M. (2022). Optimizing the airborne laser scanning estimation of basal area larger than mean (BALM): An indicator of cohort balance in forests. Ecological Indicators, 142, 109162. https://doi.org/10.1016/j.ecolind.2022.109162

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, 28(8), 2687-2704. https://doi.org/10.1111/tgis.13265

Aksoy, H., & Günlü, A. (2025). UAV and satellite-based prediction of aboveground biomass in scots pine stands: a comparative analysis of regression and neural network approaches. Earth Science Informatics, 18(1), 66. https://doi.org/10.1007/s12145-024-01657-0

Alici, N., & Dalkılıç, B. (2022). İç mekân donatı elemanlarında biyo-esaslı malzeme kullanımı. GRID-Architecture Planning and Design Journal, 5(2), 325-346. https://doi.org/10.37246/grid.941882

Anonymous, (2023). https://en-gb.topographic-map.com/map-qccstj/Mikkeli, accessed: 25.06.2025

Balazs, A., Tuominen, S., & Kangas, A. (2025). Enhancing forest inventory Accuracy: Comparing 3D-CNN and k-NN with genetic algorithm Approaches using ALS data across boreal bioregions. Computers and Electronics in Agriculture, 237, 110576. https://doi.org/10.1016/j.compag.2025.110576

Bouvier, M., Durrieu, S., Fournier, R. A., & Renaud, J. P. (2015). Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. Remote Sensing of Environment, 156, 322-334. https://doi.org/10.1016/j.rse.2014.10.004

Brown, S., Narine, L. L., & Gilbert, J. (2022). Using airborne lidar, multispectral imagery, and field inventory data to estimate basal area, volume, and aboveground biomass in heterogeneous mixed species forests: A case study in southern Alabama. Remote Sensing, 14(11), 2708. https://doi.org/10.3390/rs14112708

Cade, B. S. (1997). Comparison of tree basal area and canopy cover in habitat models: Subalpine forest. The Journal of wildlife management, 326-335.

Dube, T., Sibanda, M., Shoko, C., & Mutanga, O. (2017). Stand-volume estimation from multi-source data for coppiced and high forest Eucalyptus spp. silvicultural systems in KwaZulu-Natal, South Africa. ISPRS Journal

of Photogrammetry and Remote Sensing, 132, 162-169. https://doi.org/10.1016/j.isprsjprs.2017.09.001.

Fraser, B. T., & Congalton, R. G. (2021). Estimating primary forest attributes and rare community characteristics using unmanned aerial systems (UAS): An enrichment of conventional forest inventories. Remote Sensing, 13(15), 2971. https://doi.org/10.3390/rs13152971

Fu, L., Sharma, R. P., Zhu, G., Li, H., Hong, L., Guo, H., ... & Tang, S. (2017). A basal area increment-based approach of site productivity evaluation for multi-aged and mixed forests. Forests, 8(4), 119. https://doi.org/10.3390/f8040119

Giri, C., Ochieng, E., Tieszen, L. L., Zhu, Z., Singh, A., Loveland, T., ... & Duke, N. (2011). Status and distribution of mangrove forests of the world using earth observation satellite data. Global ecology and biogeography, 20(1), 154-159. https://doi.org/10.1111/j.1466-8238.2010.00584.x

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., ... & Townshend, J. R. (2013). High-resolution global maps of 21st-century forest cover change. science, 342(6160), 850-853. https://doi.org/10.1126/science.1244693

Hawryło, P., Francini, S., Chirici, G., Giannetti, F., Parkitna, K., Krok, G., ... & Socha, J. (2020). The use of remotely sensed data and polish NFI plots for prediction of growing stock volume using different predictive methods. Remote Sensing, 12(20), 3331. https://doi.org/10.3390/rs12203331

Hosseini, S. F., Jalilvand, H., Fallah, A., Asadi, H., & Tafazoli, M. (2025). Machine learning methods for basal area prediction of Fagus orientalis Lipsky stands based on national forest inventory. Trees, 39(2), 1-18. https://doi.org/10.1007/s00468-025-02616-y

Husch, B., Beers, T. W., & Kershaw Jr, J. A. (2002). Forest mensuration. John Wiley & Sons.

Janssen, S., Pretzsch, H., Bürgi, A., Ramstein, L., & Bont, L. G. (2023). Improving the accuracy of timber volume and basal area prediction in heterogeneously structured and mixed forests by automated co-registration of forest inventory plots and remote sensing data. Forest Ecology and Management, 532, 120795. https://doi.org/10.1016/j.foreco.2023.120795

Kangas, A., Astrup, R., Breidenbach, J., Fridman, J., Gobakken, T., Korhonen, K. T., ... & Olsson, H. (2018). Remote sensing and forest inventories in Nordic countries–roadmap for the future. Scandinavian Journal of Forest Research, 33(4), 397-412. https://doi.org/10.1080/02827581.2017.1416666

Kangas, A., Haara, A., Holopainen, M., Luoma, V., Packalen, P., Packalen, T., ... & Saarinen, N. (2019). Kaukokartoitukseen perustuvan metsävaratiedon hyötyanalyysi: Metku-hankkeen loppuraportti. https://urn.fi/URN:ISBN:978-952-326-707-7

Kelek, M. M., Cengiz, E., Oğuz, Y., & Yönetken, A. (2021). RLBP Metodu ile Mamografi Görüntülerinin İncelenmesi ve Sınıflandırılması. Afyon Kocatepe Üniversitesi Uluslararası Mühendislik Teknolojileri ve Uygulamalı Bilimler Dergisi, 4(2), 59-64.

Kelek, M. M., Fidan, U., & Oğuz, Y. (2025). The effect of wavelet transform on the classification performance of different deep learning architectures. Signal, Image and Video Processing, 19(5), 366.

Lahssini, K., Teste, F., Dayal, K. R., Durrieu, S., Ienco, D., & Monnet, J. M. (2022). Combining LiDAR metrics and Sentinel-2 imagery to estimate basal area and wood volume in complex forest environment via neural networks. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 4337-4348. https://doi.org/10.1109/JSTARS.2022.3175609

Maltamo, M., Suvanto, A., & Packalén, P. (2007). Comparison of basal area and stem frequency diameter distribution modelling using airborne laser scanner data and calibration estimation. Forest Ecology and Management, 247(1-3), 26-34.

McRoberts, R. E., Cohen, W. B., Næsset, E., Stehman, S. V., & Tomppo, E. O. (2010). Using remotely sensed data to construct and assess forest attribute maps and related spatial products. Scandinavian Journal of Forest Research, 25(4), 340-367. https://doi.org/10.1080/02827581.2010.497496

Mora, B., Wulder, M. A., White, J. C., & Hobart, G. (2013). Modeling stand height, volume, and biomass from very high spatial resolution satellite imagery and samples of airborne LIDAR. Remote Sens. 5, 2308–2326. https://doi.org/10.3390/rs5052308

Oliver, C. D., Larson, B. C., & Oliver, C. D. (1996). Forest stand dynamics (Vol. 520). New York: Wiley.

Özer, T., Akdoğan, C., Cengız, 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.

Pippuri, I., Maltamo, M., Packalen, P., & Mäkitalo, J. (2013). Predicting species-specific basal areas in urban forests using airborne laser scanning and existing stand register data. European Journal of Forest Research, 132(5), 999-1012. https://doi.org/10.1007/s10342-013-0736-8

Rodríguez, F., Pemán, J., & Aunós, Á. (2010). A reduced growth model based on stand basal area. A case for hybrid poplar plantations in northeast Spain. Forest Ecology and Management, 259(10), 2093-2102. https://doi.org/10.1016/j.foreco.2010.02.021

Ruiz-Benito, P., Madrigal-Gonzalez, J., Ratcliffe, S., Coomes, D. A., Kändler, G., Lehtonen, A., ... & Zavala, M. A. (2014). Stand structure and recent climate change constrain stand basal area change in European forests: a comparison across boreal, temperate, and Mediterranean biomes. Ecosystems, 17(8), 1439-1454. https://doi.org/10.1007/s10021-014-9806-0

Shang, C., Treitz, P., Caspersen, J., & Jones, T. (2019). Estimation of forest structural and compositional variables using ALS data and multi-seasonal satellite imagery. International Journal of Applied Earth Observation and Geoinformation, 78, 360-371. https://doi.org/10.1016/j.jag.2018.10.002

Tischendorf, W. (1943). Der Einfluss Der Exzentrizität der Schaftquerflächen Auf Das Messungsergebnis Bei Bestandesmassenermittlungen Durch Kluppung. Springer.

Toivonen, J., Korhonen, L., Kukkonen, M., Kotivuori, E., Maltamo, M., & Packalen, P. (2021). Transferability of ALS-based forest attribute models when predicting with drone-based image point cloud data. International Journal of Applied Earth Observation and Geoinformation, 103, 102484.

Véga, C., Renaud, J. P., Durrieu, S., & Bouvier, M. (2016). On the interest of penetration depth, canopy area and volume metrics to improve Lidar-based models of forest parameters. Remote Sensing of Environment, 175, 32-42. https://doi.org/10.1016/j.rse.2015.12.039

White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T., & Tompalski, P. (2016). Remote sensing technologies for enhancing forest inventories: A review. Canadian Journal of Remote Sensing, 42(5), 619-641. https://doi.org/10.1080/07038992.2016.1207484

Witzmann, S., Matitz, L., Gollob, C., Ritter, T., Kraßnitzer, R., Tockner, A., ... & Nothdurft, A. (2022). Accuracy and precision of stem cross-section modeling in 3D point clouds from TLS and caliper measurements for basal area estimation. Remote Sensing, 14(8), 1923. https://doi.org/10.3390/rs14081923

Xu, Y., Smith, S. E., Grunwald, S., Abd-Elrahman, A., & Wani, S. P. (2018). Effects of image pansharpening on soil total nitrogen prediction models in South India. Geoderma, 320, 52-66. https://doi.org/10.1016/j.geoderma.2018.01.017

Yıldız, E., Kelek, M. M., Hocaoğlu, F. O., & Oğuz, Y. (2023). Forecasting The Impact of Vaccination on Daily Cases in Turkey for Covid-19. Academic Platform Journal of Engineering and Smart Systems, 11(1), 19-26.

Youssef, A. M., Pradhan, B., Dikshit, A., Al-Katheri, M. M., Matar, S. S., & Mahdi, A. M. (2022). Landslide susceptibility mapping using CNN-1D and 2D deep learning algorithms: comparison of their performance at Asir Region, KSA. Bulletin of Engineering Geology and the Environment, 81(4), 165.

Zeng, X., Wang, D., Zhang, D., Lu, W., Li, Y., & Liu, Q. (2024). Developing the additive systems of stand basal area model for broad-leaved mixed forests. Plants, 13(13), 1758. https://doi.org/10.3390/plants13131758

Zhao, Q., Yu, S., Zhao, F., Tian, L., & Zhao, Z. (2019). Comparison of machine learning algorithms for forest parameter estimations and application for forest quality assessments. Forest Ecology and Management, 434, 224-234. https://doi.org/10.1016/j.foreco.2018.12.019

Chapter 5

Examining The Relationships Between Litterfall, Leaf Area Index, And Net Primary Productivity in Pure Anatolian Black Pine Forests

Sinan BULUT[1], Alkan GÜNLÜ[2]

Introduction

Forest primary productivity has emerged as a critical metric, given these ecosystems' integral role in the global carbon cycle and their pronounced sensitivity to climate variability. Among all components of the carbon cycle, the terrestrial pool displays the greatest unpredictability and fluctuation. Net primary production (NPP) defined as the difference between gross primary production (GPP) and autotrophic respiration represents the most dynamic element of this terrestrial carbon flux, driving year-to-year changes in land-based carbon sinks (Chapin et al., 2006; Canadell et al., 2007).

As the largest terrestrial carbon reservoirs, forests' influence on the global carbon budget is quantified by their carbon stocks, with NPP directly determining the net amount of carbon sequestered by forest stands (Field et al., 1995; Malhi et al., 2011). Variations in NPP thus serve as a vital indicator for tracking vegetation responses to climatic shifts. By capturing these changes, NPP provides a robust measure for evaluating how complex terrestrial ecosystems adapt to and are affected by climate change (Gower et al., 1999; Yu & Chen, 2016; Zhang et al., 2019).

The NPP in forest ecosystems can be measured directly in the field or inferred through either statistical or process-based models. Ground-based measurements, while accurate, often demand extensive labor and can be constrained by local environmental conditions. Statistical models employ regression analyses to link NPP (the dependent variable) with selected predictors, but their reliance on site-

[1] Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Çankırı, Türkiye, Orchid:0000-0001-6149-0910, sbulut@karatekin.edu.tr
[2] Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Çankırı, Türkiye, Orchid: 0000-0003-4759-3125, alkangunlu@karatekin.edu.tr

specific data means they lack flexibility across different temporal and spatial contexts, limiting their applicability to varying bird species. In contrast, process-based models simulate the full cascade of biological and physical processes that drive NPP. These models define a state variable for NPP alongside all direct and indirect drivers, integrate them into a cohesive framework, and then estimate NPP from first principles (Ollinger et al., 2007). Among the diverse suite of process-based tools developed for regional to global NPP estimation, remote sensing–driven models stand out. The Carnegie–Ames–Stanford Approach (CASA) model, which couples satellite data with ecosystem process simulations, is widely adopted for this purpose. In this study, we applied the CASA model to derive annual NPP estimates (Potter et al., 1993).

The leaf area index (LAI) the total one-sided leaf area per unit ground surface is a cornerstone parameter in NPP models. Since photosynthesis occurs in leaves, converting CO_2 and water into biomass, LAI serves as both a productivity indicator and a measure of canopy structure. Any disturbance—be it storm damage, snow breakage, windthrow, drought, or forestry operations—alters LAI and thus directly impacts stand productivity. High LAI values generally correspond to robust environmental productivity (Vose and Allen 1988) and show strong correlations with stand increments, growth rates, and timber volume. Long-term LAI monitoring enhances our understanding of ecosystem productivity trends and climate interactions (Kara et al., 2011; Chianucci et al., 2015; Günlü et al., 2017). Two principal methods exist for determining LAI. Direct sampling, which involves collecting leaf litter with traps and measuring leaf area per ground unit area labor intensive and time-consuming processes. Indirect optical methods, such as hemispherical photography or radiative transfer models, which leverage sensors and algorithms to estimate LAI over large areas more efficiently in terms of time and cost (Chen et al., 1997; Jonckheere et al., 2004).

Litterfall composed of leaves, branches, bark, cones, male flowers, and other materials (fruit, flowers, moss, lichen) is another vital indicator of forest health and productivity (Çakır and Akburak, 2017). Because leaves constitute the largest fraction of total litterfall, this flux represents a primary pathway for nutrients and carbon into forest soils and is a major component of NPP (Zhang et al., 2007; Shang et al., 2018).

Forest soils rank among the largest carbon reservoirs in terrestrial ecosystems, receiving roughly one-third of annual aboveground carbon inputs through litter decomposition. Consequently, decomposition rates are key predictors of soil fertility, carbon stocks, and nutrient cycling (Çömez et al., 2018; Erkan et al., 2018). They are routinely used to model soil carbon content and its temporal

dynamics as well as to validate NPP models (Evrendilek, 2006; Taşkınsu Meydan, 2008). Climatic and environmental factors heavily influence both the magnitude and seasonality of decomposition, while stand characteristics—such as canopy density, tree vigor, growth rate, basal area, age, and site productivity—further modulate decomposition processes (Erkan et al., 2018). Among the most effective means to regulate decomposition rates are targeted silvicultural interventions (Çömez et al., 2018).

Since LAI and decomposition are integral to NPP and correlate with stand metrics, elucidating their interrelationships is essential for refining forest management plans. Accurately characterizing these links can improve the timing and design of silvicultural treatments, support forest health monitoring, and ensure more consistent determination of ecosystem functions within planning units. The aim of this study is to examine and interpret the relationships between NPP, LAI, and litterfall for pure Anatolian black pine stands.

Material and Method

This study was conducted in 30 sample plots established in pure Anatolian black pine forests within the boundaries of the Yenice Forest Planning Unit in Ilgaz Forest Enterprise (Figure 1). Field measurements were made for 2020. The sample plots for litterfall measurements were established in different stands. The traps were placed in accordance with stand closure, avoiding close proximity to tree trunks and avoiding gaps within the study area. Dead cover traps were established in 30 different sample plots at different developmental stages (9-19.9 cm, 20-35.9 cm and 36<), in sites (good: site classes I and II, medium: site classes III and poor: site classes IV and V) and in closed areas (sparse: 0.10-0.40, medium: 0.41-0.70 and dense: 0.71<).

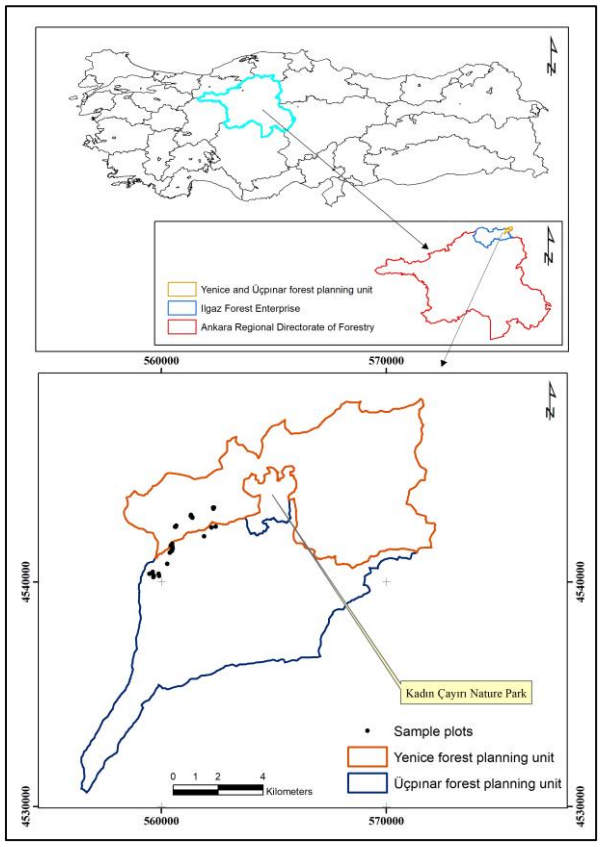

Figure 1. The study area

These established sample plots were applied to the field in a circular manner, and the following measurements and observations were made in the areas.

1. Sample plot sizes were determined between 400 and 800 m² depending on stand canopy size.
2. The geographical location and elevation of each sample plot were determined using GPS (Global Position Systems), aspect with a compass, and slope with a clinometer.
3. All trees with a diameter of 8 cm or greater at the base of the sample plot were measured with a diameter of the base of the sample plot.
4. The height and age of the tallest trees per hectare were measured using a height meter and an increment borer.
5. The age of 4-5 trees within the sample plots, close to the average diameter of the tree, was measured to determine the average age of the stand.
6. Hemispherical photographs were taken to determine the average LAI of the sample plots.

7. The litterfall amounts for each sample plot were observed from June to December (7 months).

Litterfall Sampling and Trap Design

Litterfall was collected monthly from June through December using traps positioned 1 m above ground beneath the canopy, avoiding proximity to tree trunks and gaps. Each trap featured a horizontal top to intercept falling material and a perforated base to allow water drainage (Ukonmaanaho et al., 2016). Traps were sited to sample stand-level inputs without bias from edge effects or localized canopy openings.

Laboratory Processing and Data Analysis for Litterfall

Collected litter was sealed in polyethylene bags and transported to the laboratory for sorting into needles, branches, cones, seeds, and other components (ICP-Forests 2010). After initial oven drying, subsamples were heated at 65 °C until they reached constant weight, then cooled in a desiccator and weighed to determine dry mass. To validate the CASA model, monthly dry-weight totals from each plot were statistically compared with corresponding LAI and NPP estimates.

Indirect Assessment of Leaf Area Index

In forest ecosystems, leaf area index (LAI) is a key metric for quantifying exchanges of mass and energy between vegetation and the atmosphere. As a reliable indicator of canopy structure, LAI exhibits strong links to stand growth and productivity, making it a cornerstone measurement in modern forestry studies. To derive LAI indirectly, we captured three to four hemispherical images per plot during the final days of each month. Photographs were taken at ground level facing upwards, then transferred to a digital workstation for analysis. Image processing was conducted with HemiView H. version 2.1 SR4, which computes foliage projections from canopy gaps and light transmission patterns. The LAI value for each sample plot was obtained by averaging the results from its monthly images. Due to unfavorable weather early in the study, data collection spanned seven months—from June through December 2020—ensuring consistent assessment across the study period (Figure 2).

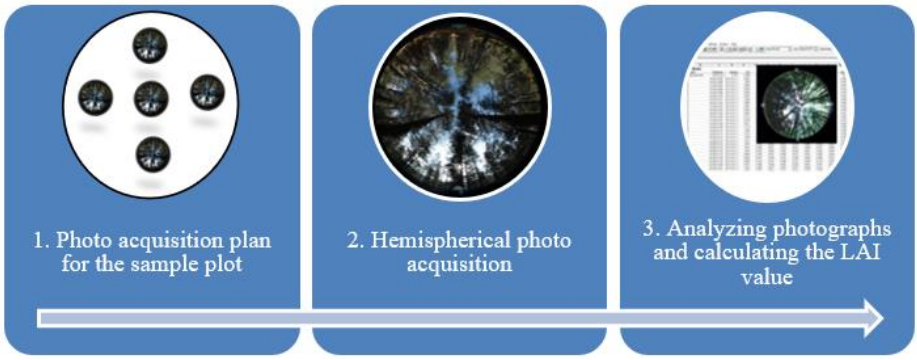

Figure 2. Calculation process of leaf area index (m^{-2} m^{-2}) values

CASA Model for Calculating Net Primary Productivity

Gross primary production (GPP) is the total amount of carbon assimilated by plants in a given region in a given time period. Net primary production (NPP) is the net amount of carbon produced by the plant through photosynthesis and remaining as a result of autotrophic respiration (AR). In other words, the net amount of carbon is taken up by the plant per unit area and time.

$$NPP = GPP - AR \quad (1)$$

Numerous models are used to determine NPP. The CASA model was used. The CASA model is a global model used to estimate net primary production (Potter, 1993). It calculates carbon flux between the atmosphere and the terrestrial biosphere at various time periods and different spatial resolutions. One of the key advantages of this model is its use of remote sensing data to mechanistically calculate net primary production and carbon sequestration using a plant and soil carbon cycle model. Therefore, it allows for the quantification of NPP, which is quite difficult, costly, and time-consuming to measure, with minimal time, expense, and effort. The CASA model can estimate NPP values using light use efficiency (ε_n) and absorbed photosynthetically active radiation (APAR) data (Cramer et al., 1999).

$$NPP = \varepsilon_n \times APAR \quad (2)$$

$$\varepsilon_n = \varepsilon_{max} \times f(T) \times f(W) \quad (3)$$

$$APAR = FPAR \times PAR \quad (4)$$

$$NPP = \varepsilon_{max} \times f(T) \times f(W) \times FPAR \times PAR \quad (5)$$

In the formula, ε_{max} represents the maximum light use efficiency, f(T) represents the temperature function, f(W) represents the precipitation function, FPAR represents the fraction of photosynthetically active radiation and PAR represents the photosynthetically active radiation (Xiao et al., 2004; Huang et al., 2010; Li and Zhou 2015; Los et al., 2000; Zhu et al. 2006).

Results and Discussion

Descriptive statistical values for LAI, litterfall, and NPP values calculated for each sample plot for June-December are given in Tables 1-3.

Table 1. Descriptive statistical values of LAI (m^{-2} m^{-2}) by months

Month	N	Minimum	Maximum	Mean	Standard Deviation
June	30	0.57	2.16	1.31	0.401
July	30	0.57	2.32	1.35	0.438
August	30	0.59	2.06	1.25	0.370
September	30	0.56	1.91	1.16	0.330
October	30	0.47	1.41	0.99	0.251
November	30	0.45	1.30	0.90	0.242
December	30	0.41	2.41	0.90	0.369

Table 2. Descriptive statistical values for total litterfall (g m^{-2} $month^{-1}$) amounts by month

Month	N	Minimum	Maximum	Mean	Standard Deviation
June	30	0.68	13.25	3.86	3.174
July	30	1.02	11.59	3.57	2.178
August	30	0.32	4.20	1.09	0.755
September	30	1.34	11.49	5.54	3.057
October	30	0.29	10.76	2.80	2.324
November	30	0.35	4.23	2.11	1.239
December	30	0.60	3.43	1.44	0.689

Table 3. Descriptive statistical values for NPP (gC m^{-2} $month^{-1}$) by month

Month	N	Minimum	Maximum	Mean	Standard Deviation
June	30	57.81	154.87	105.00	19.074
July	30	72.91	187.52	133.91	23.100
August	30	72.31	192.44	133.89	24.548
September	30	16.34	82.26	56.48	17.319
October	30	3.94	16.60	11.90	3.422
November	30	0.10	2.78	1.72	0.671
December	30	0.02	0.38	0.16	0.104

This study was conducted for pure Anatolian black pine forests located within the boundaries of the Yenice Forest Planning Unit of the Ilgaz Forest Enterprise. NPP was estimated to use the CASA model. Map of the annual total NPP was produced for pure Anatolian black pine forests located in the study area (Figure 3).

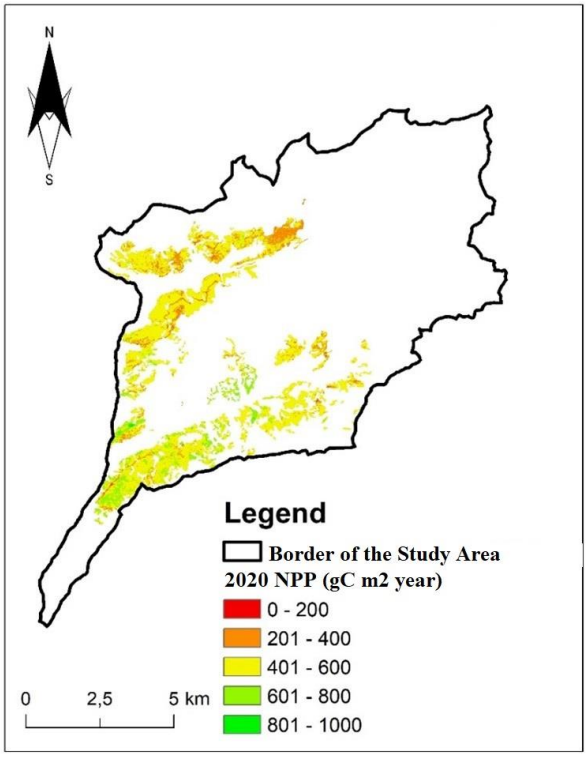

Figure 3. Map of the annual total NPP for pure Anatolian black pine forests in Yenice Forest Planning Unit

Litterfall data were obtained from 30 established sample plots between June and December 2020 (Table 4). Data could not be collected between January and May due to weather conditions. The litterfalls were divided into six classes as needle, branch, cone, bark, male flower, and other (seed, moss, lichen, mistletoe, etc.) and evaluated separately.

Table 4. Total and mean litterfall (g m^{-2} $month^{-1}$) amounts per unit area for 2020

Month	Needle	Branch	Cone	Bark	Male flower	Other	Total
June	9.767	0.860	7.103	6.467	2.171	2.564	28.933
July	14.969	1.327	4.625	2.783	2.003	2.043	27.749
August	7.464	0.153	1.070	0.243	0.065	0.055	9.051
September	39.368	3.203	0.171	1.261	1.131	0.508	45.642
October	19.184	2.244	0.667	0.390	0.435	0.413	23.333
November	15.566	0.490	0.000	0.672	0.368	0.208	17.305
December	10.942	0.259	0.000	0.094	0.157	0.078	11.530

The strength of the relationship between the NPP and litterfall data was analyzed, and the validity of the calculated NPP data was checked (Figure 4). Litterfall data from 90 buckets were used for the analysis, and average quantities for each sample plot were calculated. Although there were 90 separate litterfall data, NPP data for the area within each bucket could not be generated, so the analysis was conducted using 30 average data from the sample plots instead of the 90 data points. NPP and the litterfall components were positively correlated, with needles, cones, shells, and other litterfall. There is a negative correlation between branches and male flowers (Table 5). The highest correlation level was observed between NPP and needle litterfall, with a correlation coefficient of 0.56. Furthermore, branch and male flower litterfalls also showed correlations of 0.50 and above with NPP. A correlation level below 0.50 was also found between shells, cones, and other litterfalls and NPP. The relationship between NPP and litterfall also varies across forest types, with significant linear relationships being identified. Lv et al. (2013) reported significant relationships between NPP and litterfall in both planted and natural areas (R^2 = 0.67 / 0.30). There are strong relationships between NPP and litterfall components, and these relationship levels vary according to forest characteristics, especially forest age (Liu et al., 2019; Bulut et al., 2023a).

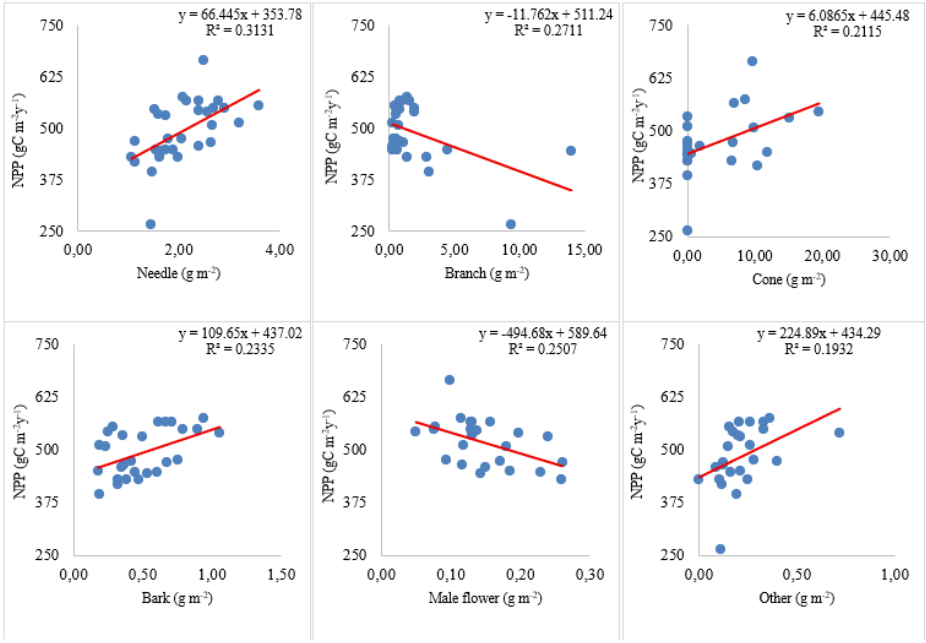

Figure 4. Examining NPP and litterfall components relationships

Table 5. Correlation levels between NPP (gC m^{-2} $year^{-1}$) and litterfall components (g m^{-2})

	Needle	Branch	Cone	Bark	Male flower	Other
r	0.56	-0.52	0.46	0.48	-0.50	0.44
R^2	0.31	0.27	0.21	0.23	0.25	0.19

The temporal relationships between LAI, NPP and litterfall were investigated. These relationships between the data obtained for the seven months from June to December were presented in scatter plots in Figure 5 and 6. It is seen that there is a positive relationship between NPP and LAI, when the temporal relationship between them is evaluated (r=0.95). NPP and LAI, which were high in June, July, and August, began to decline in September. Months with high NPP indicate that the vegetation process continues, followed by a slowdown in primary production. LAI, representing the amount of leaves per unit area and an indicator of productivity and efficiency, showed a trend consistent with NPP in this monthly observation (Bulut, 2021).

Studies have shown that LAI has a positive effect on NPP (Zhao et al., 2013). Bulut (2021) reported that there was a positive correlation between NPP and LAI at the level of 0.47. Therefore, when grouping was made according to stand parameters affecting both variables, stronger positive relationships were found.

When grouped according to stand development stages, correlation relationship levels were found to vary between 0.45 and 0.84, and when examined according to stand closure classes, it was found to vary between 0.45 and 0.62 (Bulut et al., 2023b). In addition, some studies examined the changes in both variables according to topographic parameters. In the study conducted according to elevation, it was determined that NPP decreased linearly with increasing elevation. LAI was reported to decrease quadratically with increasing elevation (Luo et al., 2004).

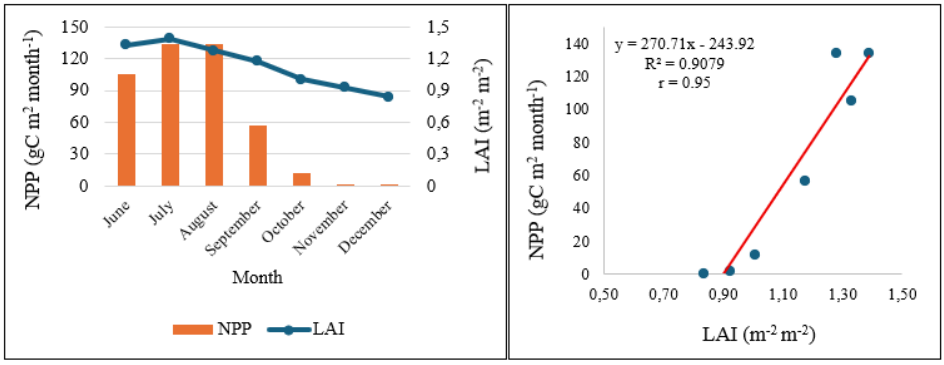

Figure 5. Relationship between monthly NPP (gC m^{-2} $month^{-1}$) and LAI (m^{-2} m^{-2})

The temporal relationship between litterfall and LAI was examined, and a correlation of 0.37 was found (Figure 6). August and September were particularly notable during the 7-month period. A more regular temporal distribution was observed for LAI. A significant decrease in litterfall occurred in August, while September determined the highest litterfall of the 7 months. These results indicate that stagnation in air mobility was observed in August, and the effects of decreasing air temperature and increasing air mobility were observed in September (Bulut, 2021). This study found a positive correlation between litterfall and LAI. Other studies have noted significant variation in LAI across forests, particularly among forests, and a negative correlation (r = -0.52) with LAI (Bloom et al., 2023). These differences may vary depending on forest characteristics, climatic conditions and environmental conditions (Erkan et al., 2018; Bulut, 2021).

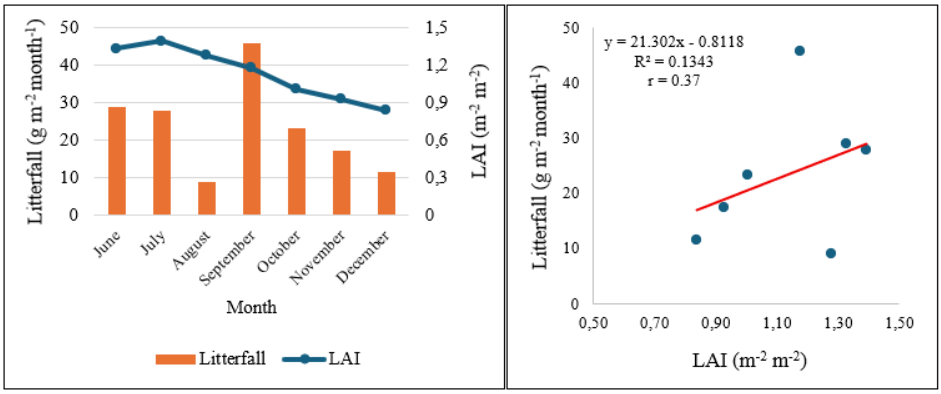

Figure 6. Relationship between monthly litterfall (g m^{-2} $month^{-1}$) and LAI (m^{-2} m^{-2})

Conclusion

NPP, LAI, and litterfall components were determined and their relationships were examined in 30 sample plots established for pure Anatolian black pine stands in the Yenice Forest Planning Unit. The field measurements were multi-temporal and conducted monthly for a seven-month period from June to December 2020. The results obtained are as follows.

- The highest correlation between NPP and litterfall components was found between needle (r = 0.54).
- A strong temporal relationship was determined between NPP and LAI (r = 0.95).
- In examining the relationships between LAI and litterfall, it was determined that the relationship levels reached higher levels, especially with the grouping according to stand parameters.

In order to achieve healthier and more stable results in these relationships, it is recommended that such studies be expanded to different species and regions and especially monitored over longer periods, and these are stated as limitations of this study.

Acknowledgements

This study was supported by the Çankırı Karatekin University, Scientific Research Projects Department, Project No: OF080120B04. Also, we thank to Mucahit Yılmaz SÖNMEZ who assisted in field measurements.

References

Bloom, D. E., Bomfim, B., Feng, Y., & Kueppers, L. M. (2023). Combining field and remote sensing data to estimate forest canopy damage and recovery following tropical cyclones across tropical regions. Environmental Research: Ecology, 2(3), 035004.

Bulut, S., Günlü, A., & Şatır, O. (2023a). Estimating net primary productivity of semi-arid Crimean pine stands using biogeochemical modelling, remote sensing, and machine learning. Ecological Informatics, 76, 102137.

Bulut, S., Günlü, A., & Keles, S. (2023b). Assessment of the interactions among net primary productivity, leaf area index and stand parameters in pure Anatolian black pine stands: A case study from Türkiye. Forest Systems, 32(1), e003-e003.

Canadell, J. G., M. U. F. Kirschbaum, W. A. Kurz, M. J. Sanz, B. Schlamadinger, and Y. Yamagata (2007), Factoring out natural and indirect human effects on terrestrial carbon sources and sinks, Environ. Sci. Policy, 10(4), 370–384.

Chapin, F. S., et al. (2006), Reconciling carbon-cycle concepts, terminology, and methods, Ecosystems, 9(7), 1041–1050.

Chen, J. M., Rich, P. M., Gower, S. T., Norman, J. M., Plummer, S. 1997. Leaf area index of boreal forests: Theory, techniques, and measurements. Journal of Geophysical Research: Atmospheres, 102(D24), 29429-29443.

Chianucci, F., Macfarlane, C., Pisek, J., Cutini, A. and Casa, R. 2015. Estimation of foliage clumping from the LAI-2000 Plant Canopy Analyzer: effect of view caps. Trees, 29(2), 355-366.

Cramer, W., Kicklighter, D.W., Bondeau, A., Iii, B. M., Churkina, G., Nemry, B., Ruimy, A., Schloss, A. L., Intercomparison, TPOTP. 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. Global Change Biology, 5, 1–15.

Çakır, M. and Akburak, S. 2017. Litterfall and nutrients return to soil in pure and mixed stands of oak and beech. Journal of The Faculty of Forestry Istanbul University, 67, 185-200.

Çömez, A., Tolunay, D. and Güner, Ş.T. 2018. Litterfall and the effects of thinning and seed cutting on carbon input into the soil in Scots pine stands in Turkey. European Journal of Forest Research, 1-14.

Erkan, N., Comez, A., Aydin, A.C., Denli, O. and Erkan, S. 2018. Litterfall in relation to stand parameters and climatic factors in Pinus brutia forests in Turkey. Scandinavian Journal of Forest Research, 33(4), 338-346.

Evrendilek, F., Berberoğlu, S., Taşkınsu Meydan, S. and Yılmaz, E. 2006. Quantifying carbon budgets of conifer Mediterranean forest ecosystems, Turkey. Environmental Monitoring and Assessment, 119(1-3), 527-543.

Field, C.B., Randerson, J.T., Malmström, C.M. 1995. Global net primary production: Combining ecology and remote sensing. Remote Sens. Environ. 51, 74–88.

Gholz, H.L. 1982. Environmental limits on aboveground net primary production, leaf area, and biomass in vegetation zones of the Pacific Northwest. Ecology, 63(2), 469-481.

Gower, S.T., Kucharik, C.J. and Norman, J.M. 1999. Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. Remote sensing of environment, 70(1), 29-51.

Günlü, A., Keleş, S., Ercanlı, İ. and Şenyurt, M. 2017. Estimation of leaf area index using WorldView-2 and Aster satellite image: a case study from Turkey. Environmental Monitoring and Assessment, 189(11), 538.

Huang, N. Niu, Z. Wu, C. and Tappert, M. C. 2010. Modeling net primary production of a fast-growing forest using a light use efficiency model. Ecological Modelling, 221(24); 2938–2948.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M. and Baret, F. 2004. Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography. Agricultural and Forest Meteorology, 121(1-2), 19-35.

Kara, Ö., Şentürk, M., Bolat, İ. ve Çakıroğlu, K. 2011. Kayın, Göknar ve Göknar-Kayın meşcerelerinde yaprak alan indeksi ile toprak özellikleri arasındaki ilişkiler. Journal of the Faculty of Forestry, Istanbul University,61(1), 47-54.

Li, Z. and Zhou, T. 2015. Optimization of forest age-dependent light-use efficiency and its implications on climate-vegetation interactions in China. In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives (Vol. 40, pp. 449–454).

Liu, X., Zhou, T., Luo, H., Xu, P., Gao, S., & Liu, J. (2019). Models ignoring spatial heterogeneities of forest age will significantly overestimate the climate effects on litterfall in China. *Science of the Total Environment*, *661*, 492-503.

Los, S. O. Pollack, N. H. Parris, M. T. Collatz, G. J. Tucker, C. J. Sellers, P. J. Dazlich, D. A. 2000. A Global 9-yr Biophysical Land Surface Dataset from NOAA AVHRR Data. Journal of Hydrometeorology, 1(2); 183–199.

Luo, T., Pan, Y., Ouyang, H., Shi, P., Luo, J., Yu, Z., & Lu, Q. (2004). Leaf area index and net primary productivity along subtropical to alpine gradients in the Tibetan Plateau. Global Ecology and Biogeography, 13(4), 345-358.

Lv, G. H., Zhou, G. S., & Wang, X. Y. (2013). Factors controlling litterfall production of forest in China. *Advanced Materials Research*, *726*, 4248-4251.

Malhi, Y., Doughty, C. and Galbraith, D. 2011. The allocation of ecosystem net primary productivity in tropical forests. Philosophical Transactions of the Royal Society B: Biological Sciences, 366(1582), 3225-3245.

Ollinger, S.V., Treuhaft, R.N., Braswell, B.H., Anderson, J.E., Martin, M.E. and Smith, M.L. 2007. The role of remote sensing in the study of terrestrial net primary production. Principles and standards for measuring primary production, 204-237.

Potter, C.S., Randerson, J.T., Field, C.B., Matson, P.A., Vitousek, P.M., Mooney, H.A. and Klooster, S.A. 1993. Terrestrial ecosystem production: A process model based on global satellite and surface data. Global Biogeochemical Cycles, 7, 811–841.

Shang, E., Xu, E., Zhang, H. and Liu, F. 2018. Analysis of spatiotemporal dynamics of the Chinese vegetation net primary productivity from the 1960s to the 2000s. Remote Sensing, 10(6), 860.

Taşkınsu Meydan, H.S. 2008. Yukarı Seyhan Havzası'nda uzaktan algılama yöntemleri ile arazi örtüsünün sınıflandırılması ve bazı orman meşcerelerinde verimliliğin modellenmesi. ÇÜ Fen Bilimleri Enstitüsü, Doktora Tezi, 194 s., Adana.

Ukonmaanaho, L., Pitman, R., Bastrup-Birk, A., Breda, N. and Rautio, P. 2016. Part XIII: Sampling and Analysis of Litterfall. In: UNECE ICP Forests Programme Co-ordinating Centre (ed.): Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Thünen Institute for Forests Ecosystems, Eberswalde, Germany, 15 p.

Vose, J.M. and Allen, H.L. 1988. Leaf area, stemwood growth, and nutrition relationships in loblolly pine. Forest Science, 34, 547-563.

Xiao, X. Hollinger, D. Aber, J. Goltz, M. Davidson, E. A. Zhang, Q. and Moore, B. 2004. Satellite-based modeling of gross primary production in an evergreen neeDÖeleaf forest. Remote Sensing of Environment, 89(4); 519–534.

Yu, B. and Chen, F. 2016. The global impact factors of net primary production in different land cover types from 2005 to 2011. SpringerPlus, 5(1), 1235.

Zhang, J., Ge, Y., Chang, J., Jiang, B., Jiang, H., Peng, C., Zhu, J., Yuan, W., Qi, L. and Yu, S. 2007. Carbon storage by ecological service forests in Zhejiang Province, subtropical China. Forest Ecology and Management, 245(1-3), 64-75.

Zhang, M., Lin, H., Wang, G., Sun, H. and Cai, Y. 2019. Estimation of vegetation productivity using a Landsat 8 time series in a heavily urbanized area, Central China. Remote Sensing, 11(2), 133.

Zhao, N., Zeng, X. F., & Zhou, J. Z. (2013). Assessment on vegetation dynamics under climate change for energy saving with satellite data and geographically weighted regression. *Advanced Materials Research*, *648*, 265-269.

Zhu, W. Pan, Y. He, H. Yu, D. and Hu, H. 2006. Simulation of maximum light use efficiency for some typical vegetation types in China. Chinese Science Bulletin, 51(4); 457–463.

Chapter 6

Predicting Individual Tree Volume by Using Multivariate Adaptive Regression Splines Technique

İlker ERCANLI[1]

Introduction

The quantitative and qualitative assessment of existing tree resources within forest ecosystems represents a pivotal element of forest planning processes. Forest inventory serves as the principal data source within this framework. By furnishing comprehensive information regarding the structural attributes and resource capacity of forest ecosystems, forest inventory plays a crucial role in sustainable forest management and operational planning procedures (Yavuz and Saraçoğlu 1999, Kapucu 2004). Within forest inventory systems, tree capital is conventionally presented as tree volume, expressed in cubic meters (m^3); forest enterprises endeavor to achieve financial equilibrium through the commercialization of timber-grade tree volume, which possesses significant economic worth. Consequently, the precise and efficient quantification of tree volume, which holds paramount importance for forest enterprise management, constitutes a fundamental prerequisite for ensuring the sustainability of forest management practices (Kalıpsız 1999, Asan et al. 2001).

Among the principal methodologies employed for tree volume determination is the development of regression models through regression analysis—a fundamental statistical technique—whereby volume is formulated as a function of diameter at breast height (DBH), a singular tree attribute amenable to facile measurement in forestry practice, in conjunction with tree height, which constitutes a critical determinant of tree volume (Kalıpsız 1999, Fırat 1973). Regression models predicting tree volume as a function of DBH and tree height have become the predominant methodology employed across diverse forestry applications, particularly within forest inventory systems, attributable to their operational practicality and demonstrable efficacy in yielding predictions of

[1] Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Çankırı, Türkiye, Orcid: https://orcid.org/0000-0003-4250-7371, ilkerercanli@karatekin.edu.tr

acceptable accuracy (Kalıpsız 1999, Kapucu 2004). The tree volume table methodology encompasses tabular representations derived from mathematical relationships between dendrometric characteristics—specifically DBH and height—and stem volume, predicated upon the estimation of tree volume through regression equations formulated via the implementation of statistical methodologies such as regression analysis (Fırat 1973, Loetsch et al. 1973). Tree volume equations constructed within the regression modeling framework, together with volume tables representing their tabulated manifestation, facilitate the prediction of stem volume through an applied approach at a specified level of precision contingent upon tree DBH and height parameters, thereby contributing substantially to the implementation of forest inventories at elevated standards of accuracy and reliability. During the 1940s, an epoch characterized by the nascent integration of computer technologies within the forestry discipline, a methodology designated as the "Graphical Method" was adopted, predicated upon the manual equilibration of relationships between tree volumes, DBH, and height on a template, utilizing graphic-based representations (Yavuz 1999, Şentürk 1997). Throughout this period, visual and manual methodologies gained widespread adoption owing to constraints in technological infrastructure. Nevertheless, concomitant with advancements in computational technologies, the deployment of computationally intensive complex regression equations became feasible, culminating in the attainment of substantially enhanced precision and accuracy in tree volume estimations.

Tree volume equations constitute regression models derived through the application of regression analysis methodology, and are capable of providing reliable, accurate, and efficient predictions contingent upon the fulfillment of various assumptions prescribed by statistical science in the implementation of said regression analysis. These assumptions include: the exhibition of normal distribution by residuals (errors), defined as the discrepancy between predicted values generated by regression equations and observational data; the manifestation of homogeneous structure in the variances of these residuals relative to the variability of predicted values (homoscedasticity condition); the absence of relationships among residuals resulting from correlation arising from temporal and spatial dependencies (the emergence of autocorrelation problems in cases of presence); and the absence of any relationships among independent variables (the materialization of multicollinearity issues in cases of presence) (Searle et al. 1992, İyit et al. 2006). As articulated by Leites and Robinson (2004), a phenomenon designated as "autocorrelation" or "serial correlation" is observed in datasets constructed through the amalgamation of diameter measurements obtained at varying heights along the stem of felled trees, which constitute the

foundation of tree volume equations. This phenomenon is characterized by the statistical interdependence of stem diameter measurements acquired from the same tree and, consequently, of prediction errors derived from these measurements. Researchers such as Searle et al. (1992) and İyit et al. (2006) characterize such correlated data structures as "hierarchical" or "clustered data" structures. The manifestation of a hierarchical and clustered data structure by sample trees selected for the construction of tree volume equations—exhibiting homogeneous characteristics at the individual level yet heterogeneous characteristics among individuals—precipitates the emergence of autocorrelation problems and engenders the calculation of confidence intervals pertaining to parameter estimates of models developed through regression analysis with systematic errors. The neglect of the influence of such data structures on predictions may constitute significant sources of error. In modeling predicated upon these interrelated hierarchically structured clustered data, Linear and Nonlinear Mixed-Effects Models, which facilitate flexible modeling of variance-covariance relationships, are being employed with increasing prevalence to incorporate dependency structures arising from repeated measurements along tree stems into statistical analyses (Keselman et al. 1998; Wolfinger & Chang 1999; Littell et al. 2005). Concurrently, Autoregressive Modeling approaches emerge as a significant alternative for mitigating autocorrelation effects manifested in measurements obtained sequentially in temporal or spatial dimensions on the same individual (Gregoire et al. 1995; Davidian & Giltinan 1995; Li & Weiskittel 2010; Li et al. 2012).

In recent years, nonparametric data mining methodologies capable of modeling complex relationships without adherence to statistical assumptions have gained prominence. Among these approaches, the Multivariate Adaptive Regression Splines (MARS) algorithm, developed in 1991 by Jerome Friedman of Stanford University, constitutes one of the significant methodologies (Friedman 1991). MARS generates a predictive model analogous to the least squares method, yet in contrast to standard regression techniques, necessitates no statistical assumptions. This characteristic renders the technique a robust alternative in circumstances where statistical assumptions cannot be satisfied. Furthermore, whether variables are continuous or discrete does not constitute a critical condition for the algorithm's applicability (Özfalcı 2008; Ünal 2009). One of the algorithm's most salient attributes is its capacity to successfully model nonlinear relationships. MARS can elucidate these complex relationships in a straightforward and efficacious manner through the utilization of piecewise functions with coefficients that vary across specific intervals of variables (Örekeci et al. 2010; Oğuz 2014). One of the most characteristic features of

MARS is its capability to successfully model nonlinear relationships between dependent and independent variables through piecewise linear functions defined with distinct coefficients according to specific value ranges of variables. In this respect, MARS is evaluated as a powerful alternative in comprehending and explicating complex interactions (Örekeci et al. 2010; Oğuz 2014).

The present study aims to investigate the potential of obtaining single-entry and double-entry tree volume predictions for forest stands through the MARS methodology. Consequently, the utilization potential and performance level of data mining techniques such as Multivariate Adaptive Regression Splines (MARS), in addition to conventional regression methods, in forestry applications will be examined in this study.

Tree Volume Calculations and Estimations

The felled sample trees obtained from forest stands are utilized for the development of single-entry (diameter at breast height) and double-entry (diameter at breast height and height) volume equations. In calculating the volume values of sample trees, each tree stem is partitioned into three fundamental sections: (i) stump, (ii) stem sections of 2-meter length, and (iii) top piece. After the volume of each section is calculated separately, the total stem volume is determined by summing these volumes.

The volume of the stump is calculated employing the cylinder formula, wherein stump base diameter ($d_{0.3}$) and height (0.3 m) values are utilized. Stem sections are partitioned into equal segments of 2-meter length, and the volume of each section is calculated using Huber's formula. In Huber's formula, the diameter at the midpoint of the section and the section length (2 m) parameters are employed. The volume of the top section is calculated utilizing the cone formula, wherein the top section base diameter (d_{top}) and height (h_{top}) values are taken into consideration. Equations (1), (2), and (3) can be employed in the volume calculations of trees. The formulae utilized in calculating tree volumes are presented below:

For the stump section; $Vstump = \frac{\pi}{4} \cdot d_{0.3}^2 \cdot 0.3$ (1)

For the top section; $Vtop = \frac{1}{3} \cdot \frac{\pi}{4} \cdot d_{top}^2 \cdot h_{top}$ (2)

Huber's formula for section volumes; $Vsections = \frac{\pi}{4} \cdot (d_{1.3}^2 + \cdots + d_n^2) \cdot 2$ (3)

In the process of developing tree volume equations, the determination phase of model parameters and associated statistical criteria is executed. While single-entry models are dependent solely on diameter at breast height, double-entry models incorporate both diameter at breast height and tree height variables. The regression analysis method was employed in the model development process. In the developed equation structures, stem volume calculated through dendrometric formulae is defined as the dependent variable, while variables more readily measurable under field conditions, such as diameter at breast height and tree height, are incorporated into the model as independent variables. For the purpose of developing regression models, various single-entry and double-entry volume equations obtained from the literature, such as those by Şentürk (1997) and Yavuz (1999), can be cited as examples. Within this context, 6 distinct model structures for single-entry volume predictions [Equations (4)-(9)] and 15 distinct model structures for double-entry volume predictions [Equations (10)-(25)] can be observed below. Regression coefficients pertaining to the models, statistical significance levels of the coefficients, and other goodness-of-fit statistics can be obtained through the SPSS 15.0 software package (SPSS Inc. 2010). Single-entry tree volume functions; $V = f(d)$

$$V = b_0 + b_1 \cdot d^2 \tag{4}$$

$$V = b_0 + b_1 \cdot d + b_2 \cdot d^2 \tag{5}$$

$$V = b_0 + b_1 \cdot d^2 + b_2 \cdot \left(\frac{1}{d}\right) \tag{6}$$

$$logV = b_0 + b_1 . logd \tag{7}$$

$$logV = b_0 + b_1 . logd + b_2 \cdot \left(\frac{1}{d}\right) \tag{8}$$

$$logV = b_0 + b_1 . logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot (logd)^4 \tag{9}$$

Double-entry tree volume functions; $V = f(d, h)$

$$V = b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot (d^2 \cdot h) \tag{10}$$

$$V = b_0 + b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot (d^2 \cdot h) \tag{11}$$

$$V = b_0 + b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot d^2 + b_4 \cdot h + b_5 \cdot (d^2 \cdot h) \quad (12)$$

$$V = b_1 \cdot d^2 + b_2 \cdot h \quad (13)$$

$$V = b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot d^2 + b_4 \cdot h^2 + b_5 \cdot (d \cdot h^2) + b_6 \cdot (d^2 \cdot h) \quad (14)$$

$$V = b_0 + b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot d^2 + b_4 \cdot h + b_5 \cdot (d \cdot h) + b_6 \cdot (d^2 \cdot h) \quad (15)$$

$$V = b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot d^2 + b_4 \cdot (d \cdot h) + b_5 \cdot (d^2 \cdot h) \quad (16)$$

$$V = b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot logh + b_4 \cdot (logh)^2 \quad (17)$$

$$V = b_0 + b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot d^2 + b_4 \cdot h^2 + b_5 \cdot (d \cdot h^2) + b_6 \cdot (d^2 \cdot h) \quad (18)$$

$$V = b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot d^2 + b_4 \cdot h + b_5 \cdot (d \cdot h) + b_6 \cdot (d^2 \cdot h) \quad (19)$$

$$V = b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot d + b_4 \cdot d^2 + b_5 \cdot (d \cdot h) + b_6 \cdot (d^2 \cdot h) \quad (20)$$

$$V = b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot d^2 + b_4 \cdot h^2 + b_5 \cdot (d \cdot h^2) + b_6 \cdot (d^2 \cdot h^2) \quad (21)$$

$$V = b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot (d \cdot h) + b_4 \cdot (d^2 \cdot h) \quad (22)$$

$$logV = b_0 + b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot logh \quad (23)$$

$$logV = b_0 + b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot logh + b_4 \cdot (logd)^4 \quad (24)$$

$$logV = b_0 + b_1 \cdot logd + b_2 \cdot \left(\frac{1}{d}\right) + b_3 \cdot (logd)^4 + b_4 \cdot logh + b_5 \cdot (logh)^4 \quad (25)$$

Where;
V : tree stem volume (m^3)
d : diameter at breast height (cm)
h : total tree height (m)
df : correction factor

log: logarithm to base 10
b_0,b_1, ..b_k: the equation coefficients.

Multivariate adaptive regression splines technique

In the estimation of tree volumes, in addition to conventional multiple regression analysis, the Multivariate Adaptive Regression Splines (MARS) method can also be employed as an alternative modeling approach. The MARS technique, proposed by Friedman (1991), constitutes a nonparametric regression approach and can operate independently of fundamental assumptions such as normality distribution, variance homogeneity, and linear relationships required by classical regression models (Özfalcı 2008). The most significant advantage of the MARS method is its capacity to model nonlinear structures by considering mutual interactions among independent variables without any functional form assumption. Consequently, it offers a powerful alternative particularly in the prediction of nonlinear and complex relationships (Örekici et al. 2005; Oğuz 2014). The operational principle of the method is predicated upon partitioning the relationship between dependent and independent variables into piecewise linear splines through basis functions. This segmentation strategy enables the modeling of complex relationships among variables with enhanced accuracy and efficiency.

As being MARS, the piecewise linear splines representing the nonlinear relationship between dependent and independent variables are connected to each other through knots of basis functions. Knots denote critical values at which the slope of the relationship changes and the transition from one linear segment to another occurs (Friedman 1991; Özfalcı 2008; Oğuz 2014). The determination of knots and the specification of linear segments between these points confer upon the MARS method the capacity to model complex nonlinear relationships in a flexible and efficient manner (Örekici et al. 2005; Oğuz 2014). Through this characteristic, an adaptive modeling approach that accommodates the local structure of the data is obtained.

The fundamental structure of MARS modeling is constituted through the connection of basis functions, created within specific value ranges of variables, at knots (Friedman 1991; Özfalcı 2008; Oğuz 2014). The mathematical formulation of this modeling approach can be defined as follows with the basis function presented in Equation (26):

$$Y = \beta_0 + \sum_{k=1}^{k} a_k \cdot (\beta_k \cdot X_t) + \varepsilon_i \quad (26)$$

In this equation, β_0 represents the constant coefficient, k denotes the number of knots, a_k represents the coefficient pertaining to the k-th basis function, $\beta_k \cdot$

X_t represents the k-th basis function for the t-th independent variable, and ε_i represents the model error.

The most significant functional components within the general equation are also (Örekici et al. 2005, Çanga and Boğa 2020):

$$\beta_k \cdot X_t = \text{maksimum}(0, \text{x} - \text{k}) = \begin{cases} x - k, x > k, \\ 0, x \leq k \end{cases} \quad (27)$$

or

$$\beta_k \cdot X_t = \text{maksimum}(0, \text{k} - \text{x}) = \begin{cases} k - x, x < k, \\ 0, x \geq k \end{cases} \quad (28)$$

can be expressed as Equation (27) and Equation (28).

The implementation process of the MARS algorithm comprises two distinct phases (Ünal 2009; Oğuz 2014; Özfalcı 2008). The forward stepwise algorithm employed in the first phase encompasses the incorporation of all potential basis functions into the system with the objective of achieving the maximum complexity level of the model. The basic functions utilized in this phase constitute mathematical structures that describe transformations of individual variables or multiple interactions among variables. Throughout the process, all variables and potential combinations thereof are analyzed stepwise, and mutual interactions among independent variables along with nonlinear transformations are integrated into the model structure. Basis function pairs are defined as mirrored basis functions where the gradient exhibits variation at the knot and the first derivative assumes a value of zero. The backward elimination algorithm or pruning process, constituting the second phase of the method, focuses on minimizing the sum of squared errors through the identification of statistically significant variables and interactions among them, consequently facilitating the attainment of the most parsimonious optimal model (Ünal 2009). This pruning process is conducted through the generalized cross-validation (GCV) technique, which takes into account both the error term and model complexity (Ünal 2009, Eyduran et al. 2019a). This process aims to identify the optimal set of basis functions that balances model complexity and generalization capability. This two-stage procedure maintains the model's generalization capability while constructing a complex and flexible model structure. The generalized cross-validation (GCV) value, which holds significance in the model development process, is calculated with Equation (29) below (Eyduran et al. 2017b, Çanga and Boğa 2020).

$$\text{GCV} = \frac{1}{n} \cdot \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{\left(1 - \frac{M + d \cdot \frac{(M-1)}{2}}{N}\right)^2} \quad (29)$$

In this formula, N represents the total number of data points, y_i denotes the observed value (volume of trees calculated by the sectional method), $\hat{y}_i$ represents the volume value of trees predicted by the MARS model, M denotes the number of basis functions, and d represents the cost-complexity measure value that can assume values between 2 and 3 (Akkuş 2017).

MARS modeling for the estimation of tree stem volumes can be implemented in the R statistical software environment (R Core Team 2024) through the "earth" package (Milborrow 2011). During the modeling process, the regularization parameter value can be maintained at the -1 level to minimize the risk of overfitting (penalty = -1). In the application of the algorithm, the k-fold cross-validation technique can be employed to enhance model reliability, and evaluation can be conducted by partitioning the dataset into 10 equal subsets (nfold = 10). This approach is recognized as a standard practice in the literature. In the model pruning phase, the GCV (Generalized Cross-Validation) criterion (pmethod = "cv") formulated in Equation (29) can be applied for the optimization of the extensive basis function pool derived in the forward stepwise procedure. In determining model complexity, lower bound (nprune = 50) and upper bound (nk = 100) values have been configured in accordance with conventional practices in the literature. In such machine learning approaches, strengthening model consistency and generalization capacity necessitates the implementation of iterative training strategies. For this purpose, utilizing the loop structure in the R language, a training procedure with a minimum of 1000 iterations can be executed for each model scenario, and the model structure possessing the highest prediction performance can be determined.

Comparison of tree volume estimations

For the purpose of comparing the performance of developed volume prediction models, ten different statistical goodness-of-fit criteria can be utilized. These criteria are as follows: (1) Mean Absolute Error (MAE), (2) Root Mean Square Error (RMSE), (3) Percent Root Mean Square Error (RMSE%), (4) Bias, (5) Percent Bias (Bias%), (6) Coefficient of Determination (R^2), (7) Total Relative Error (TRE), (8) Mean Absolute Percentage Error (MAPE), (9) Akaike Information Criterion (AIC), and (10) Bayesian Information Criterion (BIC). In evaluating model performance, when MAE, RMSE, RMSE%, Bias, Bias%, TRE, MAPE, AIC, and BIC values are as close to zero as possible, and the Coefficient of Determination (R^2) value approaches unity, the model is considered to possess high prediction accuracy. The formulae of these performance criteria utilized are presented in Equations (30)-(39).

$$MAE = \sum_{i=1}^{n} \left| V_i - \hat{V}_i \right| / n \quad (30)$$

$$RMSE = \sqrt{\sum_{i=1}^{n} \left(V_i - \hat{V}_i \right)^2 / (n-k)} \quad (31)$$

$$RMSE\% = \left(\left[\sqrt{\sum_{i=1}^{n} \left(V_i - \hat{V}_i \right)^2 / (n-k)} \right] / \bar{V}_i \right) \cdot 100 \quad (32)$$

$$Bias = \sum_{i=1}^{n} \left(V_i - \hat{V}_i \right) / n \quad (33)$$

$$Bias\% = \left(\left[\sum_{i=1}^{n} \left(V_i - \hat{V}_i \right) / n \right] / \bar{V}_i \right) \cdot 100 \quad (34)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (V_i - \hat{V}_i)}{\sum_{i=1}^{n} (V_i - \bar{V}_i)} \quad (35)$$

$$\mathrm{TRE} = \left[\frac{\sum_{i=1}^{n} \hat{V}_i - \sum_{i=1}^{n} V_i}{\sum_{i=1}^{n} V_i} \right] \cdot 100 \quad (36)$$

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{n} \left| \hat{V}_i - V_i \right|}{\sum_{i=1}^{n} V_i} \cdot 100 \quad (37)$$

$$AIC = n \cdot ln(HKOK) + 2 \cdot k \quad (38)$$

$$BIC = n \cdot ln(HKOK) + n \cdot ln(k) \quad (39)$$

In these formulas, n represents the number of data points, k denotes the number of parameters, $\hat{V}_i$ represents the volume value predicted by the volume equation, V_i represents the measured volume value, and represents the mean measured tree volume values.

In comparing the prediction performances of developed tree volume equations and models, a holistic evaluation approach encompassing all statistical goodness-of-fit criteria should be adopted. For this purpose, the relative ranking method proposed by Poudel and Cao (2013) can be applied. In traditional comparison approaches, the model exhibiting the best performance according to a specific criterion is ranked 1, followed by subsequent models ranked as 2, 3, etc. However, the relative ranking method provides a more precise and comprehensive evaluation by also considering the magnitude of performance differences among models (Özçelik and Çevlik 2017; Özdemir 2018). The

relative rank value employed in this method is calculated using Equation (40) proposed by Poudel and Cao (2013):

$$R_i = 1 + \frac{(m-1)\cdot(S_i - S_{min.})}{(S_{max.} - S_{min.})} \quad (40)$$

In this formula, R_i represents the relative rank of the i-th equation, S_i denotes the performance criterion value pertaining to the i-th equation, $S_{min.}$ represents the lowest performance criterion value, and $S_{max.}$ represents the highest performance criterion value.

Within the scope of the study, the suitability of the volume prediction model determined to be most successful through the relative ranking procedure for the stands from which the data were obtained should also be tested (model validation). The suitability analysis of volume predictions for the study area should be conducted on the dataset collected from independent sample trees not included in the model development phase. In this regard, the volumes of these independent trees should be predicted using the single-entry and double-entry equations selected as most successful, along with the MARS algorithms. In testing the suitability of predicted values, the "Paired t-Test" statistical method should be applied, and model predictions should be compared with volume values calculated through the sectional method (Kalıpsız 1988; Batu 1995). Under the condition that no statistically significant difference is observed between model predictions and volume values obtained from the sectional method ($p > 0.05$), it should be concluded that the equation or model in question is suitable for the study area. Conversely, if a statistically significant difference exists between these two value groups ($p < 0.05$), it should be concluded that the relevant equation or model is not suitable for the study area.

An exemplary application of MARS models

In this study, an exemplary application of single-entry and double-entry MARS models as an alternative approach to regression analysis in tree volume estimation has been demonstrated through MARS equations developed by Özdemir (2024). The parameter values pertaining to the single-entry MARS model generating predictions based on diameter at breast height are presented in Table 1 (Özdemir 2024). This MARS model is given in Equation (41) below (Özdemir, 2024):

$$V = 0.381527 + 0.033661 \cdot \text{Maksimum}(0, d_{1.30} - 23.0) - 0.013083 \cdot \text{Maksimum}(0, 31.0 - d_{1.30}) + 0.028978 \cdot \text{Maksimum}(0, d_{1.30} - 31.0) + 0.052703 \cdot \text{Maksimum}(0, d_{1.30} - 38.0) \quad (41)$$

In this model, represents the diameter at breast height of trees. The maximum function denotes a function that can assume either a value of 0 or a value corresponding to the difference relative to the diameter at breast height, depending on the trees' breast height diameters. To elucidate this maximum function in greater detail: for instance, when a diameter at breast height value of 20 cm is substituted into the function, the functions and are obtained, and as this function assigns the maximum value of 0 and -3 as 0, $\mathrm{Maksimum}(0{,}20-23.0)$ and $\mathrm{Maksimum}(0,-3)$, and since the coefficient 0.033661 in the model is multiplied by the value 0, this diameter value will not contribute to the calculation of volume estimation. Similarly, the function $\mathrm{Maksimum}(0,\mathrm{d}_{1.30}-31.0)$ will yield a value of 0 through the operation $\mathrm{Maksimum}(0,20-31.0)$, and the function $\mathrm{Maksimum}(0,20-38.0)$ will likewise yield a value of 0 through the operation $\mathrm{Maksimum}(0,20-38.0)$ for a diameter at breast height of 20 cm. For a diameter at breast height of 20 cm, only the function $\mathrm{Maksimum}(0,31-\mathrm{d}_{1.30})$ will assume a value of 11 through the operation Maximum(0, 31.0 - 20), and the volume estimation will be obtained by multiplying this value by the coefficient -0.013083 and subtracting the result of this multiplication from 0.381527. For a diameter at breast height of 35 cm, the functions $\mathrm{Maksimum}(0,31.0-\mathrm{d}_{1.30})$ and $\mathrm{Maksimum}(0,\mathrm{d}_{1.30}-38.0)$ will assume values of 0. Conversely, the function $\mathrm{Maksimum}(0,\mathrm{d}_{1.30}-23.0)$ fonksiyonu with $\mathrm{Maksimum}(0,35-23.0)$ will assume a value of 12 and the function Maximum(0, $d_{1.30}$ - 31.0) will assume a value of 4 through Maximum(0, 35 - 31.0), which will be multiplied by the coefficients 0.033661 and 0.028978, respectively, in the model. In this manner, volume estimations through MARS models for the values of 0 or specific values yielded by maximum functions according to relevant diameter values can be obtained for specific diameter value ranges.

The double-entry MARS model developed by Özdemir (2024), which generates predictions based on trees' diameter at breast height in conjunction with their heights, is also presented in Table 2. This MARS model is given in Equation (42) below (Özdemir, 2024):

$$V = 0.2583757 + 0.020889 \cdot \mathrm{Maksimum}(0,\mathrm{d}_{1.30}-19) - 0.008236 \cdot \mathrm{Maksimum}(0,32-\mathrm{d}_{1.30}) + 0.049773 \cdot \mathrm{Maksimum}(0,\mathrm{d}_{1.30}-32.0) + 0.04276 \cdot \mathrm{Maksimum}(0,\mathrm{d}_{1.30}-42) + 0.04276 \cdot \mathrm{Maksimum}(0,\mathrm{h}-12.0815) \quad (42)$$

In this equation, the value h represents the tree height value. The maximum functions have been explained in detail above, and particularly through the function $\mathrm{Maksimum}(0,\mathrm{h}-12.0815)$ pertaining to tree height, the result value of the

function assumes a value of 0 for the portion of tree height up to the value of 12.0815 meters. For instance, for a tree height of 10 meters, the function $\mathrm{Maksimum}(0, 10 - 12.0815)$ will assume a value of 0 through the operation $\mathrm{Maksimum}(0, -2.0815)$. Consequently, for trees lower than 12.0815 meters in height, since the maximum function will yield a value of 0 and the associated coefficient (0.042760) will be multiplied by 0, the volume estimation values of trees up to a height of 12.0815 meters with the double-entry MARS model remain constant for the same diameter values. The tree characteristic that creates the volume difference for trees up to a height of 12.0815 meters is the trees' diameter at breast height.

Table 1. Parameter values and various statistics pertaining to the single-entry MARS model (Özdemir, 2024)

Basic functions	Coefficients	Standard error	t value	p
Constant (b_0)	0.301527	0.033422	9.022	p<0.001
max(0, d_{130} - 23)	0.033661	0.005269	6.389	p<0.001
max(0, 31 - d_{130})	-0.013083	0.002324	-5.628	p<0.001
max(0, d_{130} - 31)	0.028978	0.008560	3.385	p<0.001
max(0, d_{130} - 38)	0.052703	0.007018	7.510	p<0.001

Max: maximum value

Table 2. Parameter values and various statistics pertaining to the double-entry MARS model (Özdemir, 2024)

Basic functions	Coefficients	Standard error	t value	p
Constant (b_0)	0.198376	0.043218	4.590	p<0.001
max(0, d_{130} - 19)	0.020889	0.003802	5.494	p<0.001
max(0, 32 - d_{130})	-0.008236	0.002543	-3.239	p<0.01
max(0, d_{130} - 32)	0.049773	0.004999	9.956	p<0.001
max(0, d_{130} - 42)	0.042760	0.004629	9.237	p<0.001
max(0, height-12.0815)	0.028735	0.002516	11.420	p<0.001

Max: maximum value

Discussion

In this study, concise and introductory information has been provided regarding the Multivariate Adaptive Regression Splines (MARS) technique, which constitutes an innovative method for developing single-entry and double-entry tree volume prediction equations and models, and furthermore, an exemplary application pertaining to the implementation of this method has been

elucidated. This method, the Multivariate Adaptive Regression Splines (MARS) technique—a machine learning methodology capable of operating independently of parametric assumptions and demonstrating high prediction accuracy in modeling complex nonlinear relationships—has gained prominence.

In addition to multiple regression analysis, which constitutes the classical method widely employed in tree volume estimation in forestry literature, the MARS technique, representing a nonparametric approach, can also be applied as an alternative prediction method in this research. The "earth" package written in the R programming language can be utilized in the development of MARS models.

The utilization of MARS modeling in predicting various individual tree and stand characteristics in forestry remains quite limited; Ou et al. (2019) employed MARS models alongside machine learning techniques such as Random Forest, Boosted Regression Tree, and Cubist in predicting tree diameter increment. González-Rodríguez and Diéguez-Aranda (2020) utilized the MARS model in modeling the relationships of site index with various climatic, edaphic, and physiographic variables. Guerra-Hernández et al. (2021) developed a MARS model predicting site index according to various climatic and soil properties. In our country, the utilization of MARS models has been extensively implemented in various studies within the fields of biology, economics, environment, and agriculture, whereas studies in the forestry domain remain at an initial level.

When the applicability of the MARS model is evaluated, the potential of the MARS model to yield results quite close to and, to a certain extent, superior to the outcomes pertaining to single-entry and double-entry tree volume equations obtained through multiple regression analysis regarding performance criteria is evident. In this respect, it can be concluded that the MARS model, with its capacity to provide more successful predictions compared to existing classical multiple regression equations, can be evaluated as an alternative prediction technique to the aforementioned equations. The MARS method possesses advantages such as not requiring any statistical assumptions during the model development process and being applicable to variable types including whether variables are continuous (quantitative, measurement variables) or discrete (qualitative, categorical variables) (Özfalcı 2008, Ünal 2009). Furthermore, as articulated by Orekeci et al. (2005) and Oğuz (2014), another significant characteristic of the MARS algorithm is its capability to model relationships among variables exhibiting nonlinear relationships in a straightforward manner through piecewise equations (splines) with coefficients that vary across specific intervals of variables. In this regard, MARS models possess the capability to readily and successfully model nonlinear relationships through their structures

consisting of various linear model segments. In conjunction with these advantages, the performance status of the MARS model in predicting various individual tree and stand characteristics in forestry needs to be demonstrated and evaluated.

The utilization of innovative techniques such as MARS in obtaining tree volume estimations can also facilitate the acquisition of more accurate and successful volume predictions at individual tree and stand levels, and enable forestry activities, particularly forest management, to be conducted more effectively and consistently. Specifically, the development of regional volume equations for significant forest areas of our country through recently prominent innovative techniques such as MARS can provide substantial contributions to the success of forestry activities in these areas.

References.

Akkuş, O. 2017. Hayvancılık alanında çok değişkenli uyarlanabilir regresyonun kullanımı. Yüksek Lisans Tezi, Iğdır Üniversitesi, 31 sayfa, Iğdır.

Asan, Ü. Başkent, E. Z. ve Özçelik, R. 2001. Gelişmiş ülkelerdeki ulusal orman envanter sistemleri ve Türkiye için öneriler. 1. Ulusal Ormancılık Kongresi, 19-20 Mart, 30-51, Ankara.

Batu, F. 1995. Uygulamalı İstatistik Yöntemler, KTÜ Orman Fakültesi Yayın No: 179, Trabzon, 312 s.

Çanga, D., and Boğa, M. 2020. Determination of the effect of some properties on egg yield with regression analysis method bagging MARS and R application. Turkish Journal of Agriculture - Food Science and Technology, 8(8): 1705–1712.

Davidian, M., Giltinan, D.M., 1995. Nonlinear Models for Repeated Measurement Data. New York: Chapman and Hall

Eyduran, E., Akkuş, O., Kara, K.M., Tırınk, C. and Tarıq, M. M. 2017b. Use of adaptive regression splines (MARS) in predicting body weight from body measurements in Mengali rams. International Conference On Agriculture, Forest, Sciences and Technologies. Cappadocia/Turkey, 415.

Eyduran, E., Akin, M. and Eyduran, S. P. 2019a. Application of multivariate adaptive regression splines in agricultural sciences through R Software. Nobel Bilimsel Eserler Sertifika No:20779, Ankara.

Friedman, J. H. 1991. Multivariate Adaptive Regression Splines Annals of Statistics, 19(1): 1-67.

Fırat, F. 1973. Dendrometri. IV. Baskı, İ.Ü. Orman Fakültesi, İ. Ü Yayın No, 1800, O. Yayın No, 193, Kutulmuş Matbaası, İstanbul.

González-Rodríguez, M. A. and Diéguez-Aranda, U. 2020. Exploring the use of learning techniques for relating the site index of radiata pine stands with climate, soil and physiography. Forest Ecology and Management, 458, p. 117803.

Gregoire, T. G., Schabenberger, O., Barrett, J. P. 1995. Linear modeling of irregularly spaced, unbalanced, longitudinal data from permanent plot measurement, Canadian Journal of Forest Research, 25, 137-156.

Guerra-Hernández, J., Arellano-Pérez, S., González-Ferreiro, E., Pascual, A., Altelarrea, V. S., Ruiz-González, A. D. and Álvarez-González, J. G. 2021. Developing a site index model for P. Pinaster stands in NW Spain by combining bi-temporal ALS data and environmental data. Forest Ecology and Management, 481, p. 118690.

İyit, N., Genç, A. and Arslan, F. 2006. Analysis of repeated measures for continuous response data using General Linear Model and Mixed Models, Proceedings of the international conference on modeling and simulation, Konya, TURKEY, 937-942.

Kalıpsız, A., 1988, İstatistik Yöntemler, 453 s, İ. Ü. Yayın No: 3522, İstanbul.

Kalıpsız, A. 1999, Dendrometri. İ.Ü. Orman Fakültesi Yayın No,3194/354, İstanbul.

Kapucu, F. 2004, Orman amenajmanı. KTÜ Orman Fakültesi Yay No, 215 / 33, ISBN-975-6983-35-3, 514 sayfa.

Keselman, H. J., Algina, J., Kowalchuk, R. K., Wolfinger, R. D. 1998. A Comparison of Two Approaches for Selecting Covariance Structures in the Analysis of Repeated Measures, Communications in Statistics–Computation and Simulation, 27 (3), 591–604.

Leites, L. P. and Robinson, A. P. 2004. Improving taper equations of loblolly pine with crown dimensions in a mixed-effects modeling framework, Forest Science, 50: 204-212.

Leite, H. G., Marques da Silva, M. L., Binoti, D. H. B., Fardin, L., Takizawa, F.H. 2011. Estimation of inside-bark diameter and heartwood diameter for Tectona grandis Linn. Trees using artificial neural networks, European Journal of Forest and Research 130, 263–269.

Li, R., Weiskittel, A. 2010. Development and evaluation of regional taper and volume equations for the primary conifer species in the Acadian Region, Annals of Forest Science, 67:302

Li, R., Weiskittel, A., Dick, A.R., Kershaw, J.A., Seymour, R.S., 2012. Regional stem taper equations for eleven conifer species in the Acadian region of North America: development and assessment, Northern Journal of Applied Forestry, 29:5-14.

Loetsch, F., Zöhrer, F. and Haller, K. E. 1973. Forest Inventory, Volume II, BLV Verlagsgesellschaft München Bern Wien, München, 469 p.

Milborrow, S. 2011. Derived from MDA: MARS by T. Hastie and Tibshirani, earth: Multivariate adaptive regression splines, R package.

Oğuz, A. 2014. Çok değişkenli uyarlanabilir regresyon zincirlerinin irdelenmesi ve bir uygulama, Yüksek Lisans Tezi, Erzincan Üniversitesi, 75 sayfa, Erzincan.

Ou, Q., Lei, X. and Shen, C. 2019. Individual tree diameter growth models of larch–spruce–fir mixed forests based on machine learning algorithms. Forests, 10 (2), 187.

Örekici, G., Çamdeviren, H. ve Yazıcı, C. 2005. Regresyon Modellerine Alternatif Yaklaşım: MARS, VIII. Ulusal Biyoistatistik Kongresi Bildiri Kitabı, Bursa, 105-123.

Özçelik, R. ve Çevlik, M. 2017. Batı Akdeniz yöresi doğal sedir meşcereleri için hacim denklemleri. Turkish Journal of Forestry, 18 (1): 37-48.

Örekici Temel, G., Ankaralı, H. ve Yazıcı, A. 2010. Regresyon Analizine Alternatif Bir Yaklaşım: MARS. Türkiye Klinikleri Biyoistatistik Dergisi, 2: 58-66.

Özfalcı, Y., 2008. Çok Değişkenli Uyarlanabilir Regresyon Kesitleri: MARS, Yüksek Lisans Tezi, Gazi Üniversitesi, 69 sayfa, Ankara.

Özdemir, G. 2018. Karabük yöresi kayın-göknar karışık meşcerelerinde gövde çaplarının yapay sinir ağları ile tahmin edilmesi. Yüksek Lisans Tezi, Kastamonu Üniversitesi, 101 sayfa, Kastamonu.

Özdemir, B. 2024. Ağaç hacim tahminlerinin çok değişkenli uyarlanabilir regresyon eğrileri tekniği ile elde edilmesi. Yüksek Lisans Tezi, Çankırı Karatekin Üniversitesi, 73 sayfa, Çankırı.

Poudel, K. P. and Cao, Q. V. 2013. Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. Forest Science, 59 (2): 243-252.

R Core Team, 2024. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

SPSS Institute Inc. 2010. SPSS Base 15.0, User's Guide, 750 s.

Searle, S. R., Casella, G. and Mc Culloch, C. E. 1992. Variance components, John Wiley and Sons Inc., USA.

Şentürk, N. 1997. Dişbudak (*Fraxinus angustifolia* Wahl. subps. Oxycarpa (Bieb. ExWilld.) Franco&RochaAfonso) gövde hacim ve ağaç hacim tablolarının düzenlenmesi. Yüksek lisans tezi, Karadeniz Teknik Üniversitesi, 108 sayfa, Trabzon.

Ünal, B., 2009. Çok değişkenli uyarlamalı regresyon uzanımları, Yüksek Lisans Tezi, Hacettepe Üniversitesi, 131 sayfa, İstanbul.

Wolfinger, R., Chang, M.1999. Comparing the SAS GLM and MIXED Procedures for repeated measures. SAS Institute, Inc., Cary, NC.

Yavuz, H., 1999, Taşköprü Yöresinde Karaçam İçin Hacim Fonksiyonları ve Hacim Tabloları, Turkish Journal of Agriculture and Forestry, 23, 5: 1181-1188.

Yavuz, H. ve Saraçoğlu, N. 1999. Kızılağaç için uyumlu ve uyumsuz gövde çapı modelleri. Turkish Journal of Agriculture and Forestry, 23, Ek Sayı 5: 1275-1282.

Chapter 7

Diameter Distribution Models Based on Probability Density Functions and Forestry Applications

İlker ERCANLI[1], Abdurrahman ŞAHİN[2]

Introduction

Within the framework of sustainable forestry practices, comprehensive knowledge of both the productive capacity of forest ecosystems and their structural attributes serves as an essential foundation for strategic forest management. The characterization of tree distribution patterns across diameter classifications within forest stands represents a pivotal aspect of structural analysis (Yavuz et al., 2002; Bolat, 2014). While multiple methodological frameworks and conceptualizations exist for describing stand architecture, it may be fundamentally conceptualized as the quantitative expression of how tree species are distributed among size classes in forested landscapes. Moreover, elucidating the interconnections between diameter class distribution patterns and key stand attributes—including biodiversity levels, developmental stage, stocking density, and species intermixture—holds considerable significance. Such detailed structural information constitutes an indispensable resource for numerous silvicultural applications, most notably in the context of forest management planning initiatives, thereby rendering stand structure determination a critical parameter that must inform forest-related decision-making processes.

In forest planning and economic analyses, the numerical determination and prediction of diameter distributions of trees within a stand also serve as a fundamental information source for identifying the types of products to be obtained from forests. The determination and prediction of diameter distributions of present and future forests provide an important information source for forest management and planning (Maltamo 1997). The estimation of

[1] Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Çankırı, Türkiye, Orcid: https://orcid.org/0000-0003-4250-7371, ilkerercanli@karatekin.edu.tr
[2] Artvin Çoruh University, Faculty of Forestry, Department of Forestry Engineering, Artvin, Türkiye, Orcid: https://orcid.org/0000-0002-9435-9844, asahin@artvin.edu.tr

stand diameter distributions is a critical requirement in forest planning, and in this regard, detailed analysis of stand structures enables the implementation of forestry practices in a scientific and applied manner, while also forming a foundation for sustainable forest management. Therefore, the determination of stand structures in terms of forest management and planning is of great importance for both ecosystem balance and economic benefits (Rennols et al., 1985, Borders and Patterson 1990, Laar and Akça 2007).

Probability density models characterize tree size distributions by expressing the proportional representation of stems within specific diameter categories relative to the total stand population, where proportions are constrained within a 0-1 interval (Bailey and Dell 1973). Within silvicultural research, numerous probability density models have been developed and applied for characterizing tree diameter patterns. Initial investigations by Clutter and Bennett (1965) employed a four-parameter beta formulation (Packard 2000), while subsequent methodologies—including Normal, Lognormal, Gamma, Beta, Johnson's SB, and Weibull distributions—have gained recognition as effective tools for representing stem diameter patterns in forest stands (Ercanlı and Yavuz 2010). A substantial body of literature, encompassing works by Bailey and Dell (1973), Smalley and Bailey (1974), Haffley and Schreuder (1977), Rennols et al. (1985), Knoebel et al. (1986), Pukkala et al. (1990), Saramäki (1992), Packard (2000), Liu et al. (2004), Podlaski (2006), and Palahí et al. (2007), demonstrates the application of diverse probability density approaches in modeling stem size distributions. Analysis of this literature reveals that Normal (Bailey 1980), Lognormal (Bliss and Reinker 1964), Gamma (Nelson 1964), Beta (Clutter and Bennett 1965), Johnson's SB (Johnson 1949), and Weibull formulations (Weibull 1951) have found widespread application in forestry contexts. Such methodological diversity facilitates enhanced comprehension of diameter distribution patterns and improved predictive accuracy, thereby advancing forest management strategies and promoting sustainable silvicultural practices.

Within silvicultural applications, both the three-parameter Weibull formulation and the four-parameter Johnson model have gained substantial recognition and broad adoption. The effectiveness of these probability density approaches in characterizing diverse stem size patterns stems from their parametric flexibility, accommodating either three or four adjustable coefficients. Consequently, various methodological frameworks have been developed and implemented for coefficient estimation and simulation of diameter distribution patterns. Beyond maximum likelihood estimation—the conventional approach for parametric inference in probability density models—

Lei (2008) demonstrated superior performance of the moment-based technique compared to maximum likelihood, while Liu et al. (2009) established that percentile-based estimation outperformed maximum likelihood, moment-based, and hybrid approaches. Complementing these investigations, studies conducted by Knowe (1992), Bailey et al. (1989), Knowe et al. (1997), Liu et al. (2004), Cao (2004), Poudel (2011), and Poudel and Cao (2013) have introduced methodological frameworks and mathematical expressions for determining coefficients of three-parameter Weibull or four-parameter Johnson SB formulations, utilizing moment-derived metrics including arithmetic means and standard deviations, as well as regression equations incorporating diameter measurements at the 25th, 31st, 50th, 63rd, and 95th percentiles.

This study aims to explain the importance and role of diameter distribution models, which are of considerable significance in forestry, their development processes, model structures, and the development processes of probability density functions that are widely used in developing these models. In this context, summary information will be provided regarding the processes related to the development of probability density functions, and the usage status of the model structure will also be demonstrated with examples.

Probability Density Functions

Within silvicultural research, "probability density functions" (pdf) constitute essential statistical instruments employed for characterizing stem diameter patterns and projecting tree populations across size classifications or intervals. These mathematical constructs facilitate comprehensive depiction of stand architecture by projecting stem quantities within designated diameter categories across specified size gradients. Probability density models serve as analytical frameworks that quantify the proportional representation of individuals (f, frequency) occurring at a particular diameter measurement (x) or within a size interval relative to the overall stand population, yielding values constrained between 0 and 1 (Bailey and Dell 1973). Furthermore, these formulations enable computation of cumulative proportions, representing the aggregated fraction of individuals from the lower bound up to a specified diameter threshold relative to total stand composition. Such models estimate the proportional contribution of observations at any given diameter input value relative to the complete tree population within that stand, with predicted values necessarily bounded within the 0-1 interval. Through this analytical approach, stem quantities at discrete diameter measurements or across diameter intervals can be projected using probability density modeling techniques.

In forestry literature, various studies have been conducted on the comparison of different probability density functions. In these studies, particularly due to their flexibility and ability to represent different distributions, the 3-parameter Weibull and 4-parameter Johnson SB probability density functions stand out as methods that achieve successful results in modeling diameter distributions (Smalley and Bailey 1974, Lohrey and Bailey 1977, Feduccia et al., 1979, Matney and Sullivan 1982, Clutter et al., 1984, Baldwin and Feduccia 1987, Liu et al., 2004, Čavlović et al., 2006, Palahí et al., 2006, Zhang and Liu 2006, Gorgoso-Varela et al., 2007, Jiang and Brooks 2009, Andrašev et al., 2009). In this study, the 3-parameter Weibull and 4-parameter Johnson SB probability density functions, which have previously achieved successful results in modeling different stand structures, were preferred for estimating the distributions of trees across diameter classes and developing diameter distribution models. The formulas for these distribution functions are given in Equations (1) and (2):

3-parameter Weibull probability density function (Bailey and Dell 1973):

$$F(x,\alpha,\beta,\gamma) = \frac{\gamma}{\beta}\cdot\left(\frac{x-a}{\beta}\right)^{\gamma-1}\cdot exp\left(-\left(\frac{x-a}{\beta}\right)^{\gamma}\right) \quad (1)$$

Where;
x: diameter value,
α: location parameter,
β: scale parameter,
γ: shape parameter.

4-parameter Johnson SB Probability Density Function (Johnson 1949):

$$F(x,\delta,\lambda,\gamma,\xi) = \frac{\delta}{\lambda\cdot\sqrt{2\cdot\pi}\cdot z\cdot(1-z)}\cdot exp\left(-\frac{1}{2}\cdot\left(\gamma+\delta\cdot ln\left(\frac{z}{1-z}\right)\right)^2\right) \quad (2)$$

$$z = \frac{x-\xi}{\lambda}$$

Where;
x: çap değeri
λ: scale parameter
ξ: location parameter as the lower limit value
γ: first shape parameter
δ: second shape parameter

Estimation of Probability Density Function Parameters

Various methods can be used for estimating the parameters of the 3-parameter Weibull and 4-parameter Johnson SB probability density functions, including Nonlinear Regression Analysis, Maximum Likelihood Estimation, Moment-Based Parameter Recovery equations, and Percentile-Based Parameter Recovery equations (Bolat 2014).

Within silvicultural applications, parametric estimation for both three-parameter Weibull and four-parameter Johnson SB formulations via maximum likelihood inference is accomplished using the "optim" package implemented in R statistical computing environment (R Core Team 2024). This computational tool employs the Nelder-Mead simplex algorithm (Nelder and Mead 1965), recognized for its rapid convergence and computational efficiency (Ogana et al. 2020). Furthermore, during the parametric estimation process for probability density models, stem frequencies across diameter categories undergo standardization to per-hectare metrics, with appropriate weighting applied to account for variability in sampling unit dimensions.

Regarding the equations based on moments and percentiles of diameter distributions, which are other methods used in estimating the parameters of probability density functions, the studies of Cao (2004) and Poudel and Cao (2013) stand out for the 3-parameter Weibull function. In these studies, while the equations based on moments of the diameter distribution use the minimum value of diameter values, variance value, arithmetic mean ($\bar{d}$), and median diameter ($\tilde{d}$) calculated as the diameter of the tree with mean basal area; in the equations based on percentiles, diameter values corresponding to the 31st, 50th, 63rd, and 95th percentile ranks are used when diameter values are ordered from smallest to largest. In the equations referred to as hybrid, it is observed that there is an approach in which the arithmetic means and median diameter values of the diameter values are used together with the 25th, 31st, 50th, 63rd, and 95th percentile values. The methods used to estimate the parameters of the 3-parameter Weibull function can be expressed as follows: Method (1): maximum likelihood method based on the Nelder-Mead optimization algorithm, Method (2): equations using the variance value and arithmetic mean ($\overline{D}$) values of diameters, which are moment values of the diameter distribution, Method (3): equations using other moment values such as the arithmetic mean ($\overline{D}$) and quadratic mean diameter (D_g) calculated as the diameter of the tree with mean basal area, Method (4): equations using diameter values corresponding to the 31st and 63rd percentile values when diameter values are ordered from smallest to largest, Method (5): equations using diameter values corresponding to the

50th and 95th percentile values as another application of the percentile method, Method (6): equations using the quadratic mean diameter (D_g) and diameter values corresponding to the 25th and 95th percentile values as a method referred to as the hybrid method (Poudel and Cao 2013) that includes both moment and percentile values of the diameter distribution, Method (7): equations using the quadratic mean diameter (D_g) and diameter values corresponding to the 31st and 63rd percentile values as another application of the hybrid method, and Method (8): equations using the quadratic mean diameter (D_g) and diameter values corresponding to the 25th, 50th, and 95th percentile values as another application of the hybrid method—eight different methods are used. The formulas for these methods related to the 3-parameter Weibull function expressed in the studies of Cao (2004) and Poudel and Cao (2013) are presented in Table 1. In applying these methods for the 3-parameter Weibull function, the "tidyverse", "mass", and "actuar" libraries are used and relevant coding can be performed.

For parametric inference of the Johnson SB formulation utilizing moment-based equations and diameter distribution percentiles, several methodological approaches detailed by Scolforo et al. (2003) and Gorgoso et al. (2012) are implemented, including Conditional Maximum Likelihood (CML), the Knoebel and Burkhart (1991) technique, Mode-based estimation, and Moment-based estimation. Coefficient determination for the four-parameter Johnson SB model through these approaches incorporates minimum and maximum diameter observations alongside the 50th and 95th percentile thresholds. Five distinct estimation frameworks are employed for deriving coefficients of the four-parameter Johnson SB formulation: Approach (1) implements maximum likelihood estimation via the Nelder-Mead simplex algorithm; Approach (2) applies Conditional Maximum Likelihood estimation; Approach (3) utilizes the Knoebel and Burkhart methodology; Approach (4) employs Mode-based estimation; and Approach (5) adopts Moment-based estimation. Mathematical expressions for these estimation procedures concerning the four-parameter Johnson SB model, as documented in Scolforo et al. (2003) and Gorgoso et al. (2012), are compiled in Table 2. Computational implementation of these methodologies for the Johnson SB formulation requires the "dplyr", "tidyr", "SuppDists", "moments", and "optimx" packages within the R programming environment, facilitating appropriate algorithmic execution.

Table 1. Equations for estimating 3-parameter Weibull function parameters based on moment and percentile methods (Cao 2004 and Poudel and Cao 2013)

Parameter estimation methods	Equations
Method (1): maximum likelihood method based on the Nelder-Mead optimization algorithm	$\frac{\sum_{i=1}^{n}(x_i)^{\gamma} ln(x_i)}{\sum_{i=1}^{n}(x_i)^{\gamma}} - \frac{1}{\gamma} = \frac{1}{n}\sum_{i=1}^{n}(x_i)$ $\beta = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i)^{\gamma}\right)^{\frac{1}{c}}$
Method (2): equations using the variance value and arithmetic mean ($\overline{D}$) values of diameters	$\alpha = 0.5 \cdot d_{min} \quad \beta = \frac{\overline{D}-\alpha}{G_1}$ $\gamma: \beta^2(G_2 - G_1^2) - d_{var} = 0$
Method (3): equations using other moment values such as the arithmetic mean ($\overline{D}$) and quadratic mean diameter (D_g) calculated as the diameter of the tree with mean basal area	$\alpha = 0.5 \cdot d_{min}$ $\beta = -\alpha \cdot \left(\frac{G_1}{G_2}\right) + \left[\left(\frac{\alpha}{G_2}\right)^2 \cdot (G_1^2 - G_2) + \frac{D_g^2}{G_2}\right]^{0.5}$ $\gamma: \beta^2(G_2 - G_1^2) - d_{var} = 0$
Method (4): equations using diameter values corresponding to the 31st and 63rd percentile values when diameter values are ordered from smallest to largest	$\alpha = 0.5 \cdot d_{min} \quad \gamma = \frac{Ln\left(\frac{Ln(1-0.63)}{Ln(1-0,31)}\right)}{Ln(d_{\%63}-\alpha)-Ln(d_{\%31}-\alpha)}$ $\beta = \frac{d_{\%63} - \alpha}{\left(-Ln(1-0.63)\right)^{\frac{1}{\gamma}}}$
Method (5): equations using diameter values corresponding to the 50th and 95th percentile values as another application of the percentile method	$\alpha = 0.5 \cdot d_{min} \quad \gamma = \frac{Ln\left(\frac{Ln(1-0.95)}{Ln(1-0,50)}\right)}{Ln(d_{\%95}-\alpha)-Ln(d_{\%50}-\alpha)}$ $\beta = \frac{d_{\%50} - \alpha}{\left(-Ln(1-0.50)\right)^{\frac{1}{\gamma}}}$
Method (6): equations using the quadratic mean	$\alpha = 0.5 \cdot d_{min} \quad \gamma = \frac{Ln\left(\frac{Ln(1-0.95)}{Ln(1-0,25)}\right)}{Ln(d_{\%95}-\alpha)-Ln(d_{\%25}-\alpha)}$

diameter (D_g) and diameter values corresponding to the 25th and 95th percentile values	$\beta = -\alpha \cdot \left(\frac{G_1}{G_2}\right) + \left[\left(\frac{\alpha}{G_2}\right)^2 \cdot (G_1^2 - G_2) + \frac{D_g^2}{G_2}\right]^{0.5}$
Method (7): equations using the quadratic mean diameter (D_g) and diameter values corresponding to the 31st and 63rd percentile values	$\alpha = 0.5 \cdot d_{min} \quad \gamma = \frac{Ln\left(\frac{Ln(1-0.63)}{Ln(1-0,31)}\right)}{Ln(d_{\%63}-\alpha)-Ln(d_{\%31}-\alpha)}$ $\beta = -\alpha \cdot \left(\frac{G_1}{G_2}\right) + \left[\left(\frac{\alpha}{G_2}\right)^2 \cdot (G_1^2 - G_2) + \frac{D_g^2}{G_2}\right]^{0.5}$
Method (8): equations using the quadratic mean diameter (D_g) and diameter values corresponding to the 25th, 50th, and 95th percentile	$\alpha = \frac{n^{0.3333} \cdot d_{min} - d_{\%50}}{n^{0.3333} - 1}$ $\gamma = \frac{Ln\left(\frac{Ln(1-0.95)}{Ln(1-0,50)}\right)}{Ln(d_{\%95} - \alpha) - Ln(d_{\%50} - \alpha)}$ $\beta = -\alpha \cdot \left(\frac{G_1}{G_2}\right) + \left[\left(\frac{\alpha}{G_2}\right)^2 \cdot (G_1^2 - G_2) + \frac{D_g^2}{G_2}\right]^{0.5}$

In the equations; n: number of trees in the sample plot, d_{min}: minimum diameter in the sample plot, a: minimum diameter, s^2: variance of diameter values, d_{25}, d_{31}, d_{50}, d_{63}, and d_{95}: diameter values corresponding to the 25th, 31st, 50th, 63rd, and 95th percentile data when the data in the diameter distribution are ordered from smallest to largest, Dg: Quadratic mean diameter: diameter of the tree with mean basal area calculated for the sample plot, and $G_1$and G_2: values corresponding to the Gamma function.

Table 2. Equations for estimating 4-parameter Johnson SB function parameters (Scolforo et al., 2003 and Gorgoso et al., 2012)

Parameter estimation methods	**Equations**
Method (1): maximum likelihood method based on the Nelder-Mead optimization algorithm	ξ: $\%d_{min}$ λ: $d_{max} - d_{min}$ δ: $\frac{1}{S_f}$ γ: $\frac{-\bar{f}}{S_f}$
Method (2): Conditional Maximum Likelihood Method	ξ: $d_{min} - 1.3$ λ: $d_{max} - \xi + 3.8$ δ: $\frac{1}{S_f}$ γ: $\frac{-\bar{f}}{S_f}$
Method (3): Knoebel and Burkhart Method	ξ: $d_{min} - 1.3$ λ: $d_{max} - \xi + 3.8$ δ: $\frac{Z_{95}}{\ln\left(\frac{d_{\%95} - \xi}{\xi + \lambda - d_{\%95}}\right) - \ln\left(\frac{d_{\%50} - \xi}{\xi + \lambda - d_{\%50}}\right)}$ γ: $-\delta \cdot \left(\frac{d_{\%50} - \xi}{\xi + \lambda - d_{\%50}}\right)$
Method (4): Mode Method	ξ: $d_{min} - 1.3$ λ: $d_{max} - \xi + 3.8$ δ: $\frac{4}{4 \cdot \sigma_x}$ γ: $\frac{2 \cdot x_m - 2 \cdot \xi - \lambda}{\lambda \cdot \delta} - \delta \cdot ln\left(\frac{x_m - \xi}{\lambda + \xi - x_m}\right)$
Method (5): Moments Method	ξ: $\%d_{min}$ λ: $d_{max} - d_{min}$ δ: $\frac{\mu(1-\mu)}{Sd(x)} + \frac{Sd(x)}{4}\left[\frac{1}{\mu(1-\mu)} - 8\right]$ γ: $\delta \cdot ln\left(\frac{1-\mu}{\mu}\right) + \left(\frac{0.5-\mu}{\delta}\right)$

In the equations, mean of function $\bar{f}$; $f = ln[(d_i - \xi)/(\xi - \lambda - d_i)]$, d_i; midpoint value of the diameter class to which trees belong, S_f; standard deviation of the f function, μ value; $\mu = \left(\bar{d} - \xi\right)/\ \lambda$ and, $\bar{d}$: mean diameter, $Sd(x)$; adjusted standard deviation, with the formula: $\sigma_x/\ \lambda$'dir, σ_x; standard deviation of diameter in sample plots, $d_{\%25}$, $d_{\%31}$, $d_{\%50}$ $d_{\%63}$, $d_{\%95}$; diameter values corresponding to the 25th, 31st, 50th, 63rd, and 95th percentile data when the data in the diameter distribution are ordered from smallest to largest, Z_{95}; $d_{\%95}$ value corresponding to the p value from the standard normal distribution table, x_m; mode value of diameter in the sample plot, d_{max}; maximum diameter value in the sample plot, and d_{min}: minimum diameter value in the sample plot.

Discussion

This investigation presents and elucidates multiple parametric estimation techniques and mathematical formulations for three-parameter Weibull and four-parameter Johnson SB models, which represent commonly employed probability density approaches in stem diameter distribution modeling research. Coefficient determination for these formulations incorporates maximum likelihood inference, diverse percentile-derived and moment-derived metrics from diameter distributions, as well as composite methodologies integrating both statistical moments and percentile thresholds. Literature review reveals seminal contributions by Cao (2004) and Poudel and Cao (2013) regarding three-parameter Weibull formulations, alongside notable work by Scolforo et al. (2003) and Gorgoso et al. (2012) concerning four-parameter Johnson SB specifications. For the three-parameter Weibull model, eight alternative estimation frameworks are implemented, encompassing maximum likelihood inference, moment-based techniques derived from diameter distributions, percentile-based approaches, and integrated methodologies combining moment and percentile statistics. Specifically, eight distinct computational strategies are applied for three-parameter Weibull coefficient estimation. The first strategy implements maximum likelihood inference via the Nelder-Mead optimization algorithm. Method 2 incorporates equations that employ the variance and arithmetic mean ($\bar{d}$) of diameter values, representing moment-based parameters of the diameter distribution. Method 3 applies equations utilizing additional moment-based parameters, specifically the arithmetic mean ($\bar{d}$) and the quadratic mean diameter (Dg), the latter being computed as the diameter of the tree with mean basal area. Methods 4 and 5 represent percentile-based approaches: Method 4 employs equations based on diameter values

corresponding to the 31st and 63rd percentiles of the ordered diameter distribution, while Method 5 utilizes diameter values at the 50th and 95th percentiles. Methods 6, 7, and 8 constitute hybrid approaches that integrate both moment and percentile values of the diameter distribution (Poudel and Cao 2013). Specifically, Method 6 combines the quadratic mean diameter (Dg) with diameter values at the 25th and 95th percentiles, Method 7 incorporates the quadratic mean diameter (Dg) with values at the 31st and 63rd percentiles, and Method 8 integrates the quadratic mean diameter (Dg) with diameter values corresponding to the 25th, 50th, and 95th percentiles. In calculating the 4-parameter Johnson SB function, equations related to five different methods are used: Maximum Likelihood Method based on the Nelder-Mead optimization algorithm (Method 1), Conditional Maximum Likelihood Method (Method 2), Knoebel and Burkhart Method (Method 3), Mode Method (Method 4), and Moments Method (Method 5).

Review of antecedent investigations on this subject reveals diverse percentile-based methodologies for Weibull coefficient estimation in diameter distribution modeling. Lohrey and Bailey (1977) incorporated diameter measurements at the 24th and 93rd percentile thresholds for parametric inference of the Weibull formulation. McTague and Bailey (1987) derived Weibull coefficients utilizing diameter observations at the 10th, 63rd, and 93rd percentiles. Bailey et al. (1989) conducted parametric estimation for the Weibull model by integrating minimum diameter, quadratic mean diameter (Dg), and measurements at the 25th, 50th, and 95th percentiles. Bullock and Burkhart (2005) determined Weibull coefficients through quadratic mean diameter (Dg) combined with 25th and 97th percentile thresholds. Comparative analyses assessing predictive efficacy across various percentile-based approaches by Poudel (2011) and Karakaş (2013) demonstrated optimal performance when employing quadratic mean diameter (dg) alongside 25th, 50th, and 95th percentile measurements. Bolat (2014) recognized the methodology incorporating 31st and 63rd percentile thresholds as demonstrating superior efficacy in characterizing diameter distributions across heterogeneous species compositions and stand configurations. Analogously, multiple investigations utilizing alternative percentile combinations for Weibull parametrization (Brooks et al., 1992; Knowe et al., 2005; Lee and Coble, 2006; Coble and Lee, 2008; Jiang and Brooks, 2009) have exhibited satisfactory modeling outcomes for diameter distributions. Ercanlı et al. (2013) computed three-parameter Weibull coefficients by integrating selected percentile and moment statistics from diameter distributions, concluding that the framework employing 25th, 50th, and 63rd percentile thresholds yielded superior predictive

performance. Subsequently, Ercanlı et al. (2016) verified that equation-based parametric estimation utilizing 31st, 50th, and 63rd percentile measurements constituted the most efficacious methodology for characterizing stand diameter distributions within their investigation area.

Models characterizing diameter distributions, stand dynamics, and productivity tables facilitate enhanced precision in estimating stand architectural attributes within silvicultural operations. Implementation of these analytical frameworks advances forest management strategies and sustainability objectives by generating comprehensive projections of stand volume, basal area, and stem frequency across diameter classifications. Within the national context, Normal Yield Tables and Density-Dependent Yield Tables serve as primary instruments for projecting increment and development patterns. Nevertheless, existing inventory methodologies exhibit limitations in providing sufficiently granular structural characterizations, necessitating more refined predictive capabilities. Modeling frameworks constructed upon probability density formulations address these deficiencies by enabling inventory projection at the diameter class resolution. Specifically, projections derived through integration of diverse probability density formulations with Density-Dependent Yield Tables permit estimation of distribution patterns for attributes including stem frequency, basal area, and volume across stand diameter classifications. This synthetic methodology enables acquisition of essential fine-scale estimates requisite for forest management and silvicultural planning initiatives.

When characterizing stem diameter patterns across forests exhibiting heterogeneous stand architectures, accounting for species diversity and site-specific characteristics becomes imperative. Additionally, recognizing diverse coefficient estimation methodologies proves fundamental for robust development of diameter distribution modeling frameworks. Projecting future diameter patterns necessitates establishing semi-permanent or permanent monitoring installations coupled with systematic temporal measurements, representing critical prerequisites. Accordingly, implementing diameter distribution models alongside semi-permanent monitoring installations constitutes indispensable instruments for optimizing planning efficacy and operational management within silvicultural contexts. Such methodologies prove vital for maintaining sustainability and operational efficiency throughout the forestry sector. The diameter distribution frameworks constructed within this investigation utilized data derived from temporary sampling installations; nevertheless, achieving superior elucidation of variability across individual tree attributes and stand characteristics requires establishing permanent or semi-permanent experimental installations where stands undergo continuous

monitoring across designated temporal intervals with systematic measurement protocols. Ultimately, instituting permanent and semi-permanent experimental installations accompanied by systematic measurement regimens represents a critical imperative for advancing forestry practice within the national context.

References

Andrašev, S., Bobinac, M., & Orlović, S. (2009). Diameter structure models of Black Poplar selected clones in the section Aigeiros (Duby) obtained by the Weibull distribution. *Šumarski list*, *133*(11-12), 589-603.

Bailey, R. L., & Dell, T. R. (1973). Quantifying diameter distributions with the Weibull function. Forest Sci 19(2), 97- 104.

Bailey, R. L., Burgan, T. M., & Jokela, E. J. (1989). Fertilized midrotation-aged slash pine plantations—stand structure and yield prediction models. *Southern Journal of Applied Forestry*, *13*(2), 76-80.

Baldwin, V.C., Jr., & Feduccia, D.P. (1987). *Loblolly pine growth and yield prediction for managed West Gulf plantations*. USDA Forest Service, Southern Forest Experiment Station, Research Paper SO-236. n. 27 p.

Bliss, C. I., & Reinker, K. A. (1964). A lognormal approach to diameter distributions in even-aged stands. *Forest Science*, *10*(3), 350-360.

Bolat, F. (2014). Developing some diameter distribution models for stands located Kestel forest unit directorate, Bursa Forest district directorate. Çankırı Karatekin Universiy, Çankırı, 85 p.

Borders, B. E., & Patterson, W. D. (1990). Projecting stand tables: a comparison of the Weibull diameter distribution method, a percentile-based projection method, and a basal area growth projection method. *Forest Science*, *36*(2), 413-424.

Brooks, J. R., Borders, B. E., & Bailey, R. L. (1992). Predicting diameter distributions for site-prepared loblolly and slash pine plantations. *Southern Journal of Applied Forestry*, *16*(3), 130-133.

Bullock, B. P., & Burkhart, H. E. (2005). Juvenile diameter distributions of loblolly pine characterized by the two-parameter Weibull function. *New Forests*, *29*(3), 233-244.

Cao, Q. V. (2004). Predicting parameters of a Weibull function for modeling diameter distribution. *Forest science*, *50*(5), 682-685.

Čavlović, J., Božić, M., & Boncina, A. (2006). Stand structure of an uneven-aged fir–beech forest with an irregular diameter structure: modeling the development of the Belevine forest, Croatia. *European Journal of Forest Research*, *125*(4), 325-333.

Clutter, J. L., & Bennett, F. A. (1965). Diameter distributions in old-field slash pine plantations. *Georgia Forest Research Council, Report No. 13*: 9p. USA.

Clutter, D. J., Leopold, D. A., & Gould, L. V. (1984). Stand structure and yields of site-prepared loblolly pine plantations in the lower coastal plain of the

Carolinas, Georgia, and North Florida. USDA Forest Service, Gen. Tech. Rep. SE-27: 173 p.

Coble, D. W., & Lee, Y. J. (2008). A new diameter distribution model for unmanaged slash pine plantations in East Texas. *Southern Journal of Applied Forestry*, *32*(2), 89-94.

Ercanlı İ, & Yavuz H. (2010). The probability density functions to diameter distributions for oriental spruce and Scots pine mixed stands. Kastamonu Univ J Forest Fac 10(1), 68-83.

Ercanlı, İ., Bolat, F., & Kahriman, A. (2013). Comparing parameter recovery methods for diameter distribution models of Oriental spruce (*Picea orientalis* (L.) Link.) and Scotch pine (*Pinus sylvestris* L.) mixed stands located Trabzon and Giresun Forest Regional Directorate. International Caucasian Forestry Symposium (24-26 October), pp. 119-126, Artvin-Türkiye.

Ercanlı, İ., Bolat, F., & Kahriman, A. (2016). Modelling stand diameter distribution by using 3-parameters Weibull probability density function Sarıçiçek-Vezirköprü Forest Enterprise. *Anatolian Journal of Forest Research*, 2(1-2), 13-24.

Feduccia, D. P., Dell, T. R., Mann Jr., W. F., Campbell, T. E., & Polmer, B. H. (1979). Yields of unthinned loblolly pine plantations on cutover sites in the West Gulf region. USDA Forest Service, Research Paper SO -148, 88 p.

Gorgoso-Varela, J. J., Gonzalez, J. Á., Rojo, A. & Grandas-Arias, J. A. (2007). Modelling diameter distributions of *Betula alba* L. stands in northwest Spain with the two-parameter Weibull function. *Forest Systems*, *16*(2), 113-123.

Gorgoso, J. J., Rojo, A., Cámara-Obregón, A., & Diéguez-Aranda, U. (2012). A comparison of estimation methods for fitting Weibull, Johnson's SB and beta functions to Pinus pinaster, Pinus radiate and Pinus sylvestris stands in northwest Spain. *Forest systems*, *21*(3), 446-459.

Hafley, W. L., & Schreuder, H. T. (1977). Statistical distributions for fitting diameter and height data in even-aged stands. *Canadian Journal of Forest Research*, *7*(3), 481-487.

Jiang, L., & Brooks, J. R. (2009). Predicting diameter distributions for young longleaf pine plantations in Southwest Georgia. *Southern Journal of Applied Forestry*, *33*(1), 25-28.

Johnson, N. L. (1949). System of frequency curves generated by methods of translation. *Biometrika*, 36(1-2), 149-176.

Karakaş, R. (2013). Modeling diameter distributions in önsen natural stone pine (*Pinus pinea* L.) stands. Kahramanmaraş Sütcü İmam University, Kahramanmaraş 84 p.

Knoebel, B. R., & Burkhart, H. E. (1991). A bivariate distribution approach to modeling forest diameter distributions at two points in time. *Biometrics*, 241-253.

Knoebel, B. R., Burkhart, H. E., & Beck., D. E., (1986). A growth and yield model for thinned stands of yellow-poplar. Forest Science, 32, Issue suppl_2, June 1986, pages a0001-z0002,

Knowe, S. A. (1992). Predicting the impact of interspecific competition in young loblolly pine plantations with diameter distribution models. *Forest ecology and management*, *55*(1-4), 65-82.

Knowe, S. A., Ahrens, G. R., & DeBell, D. S. (1997). Comparison of diameter-distribution-prediction, stand-table-projection, and individual-tree-growth modeling approaches for young red alder plantations. *Forest Ecology and Management*, *98*(1), 49-60.

Knowe, S. A., Radosevich, S. R., & Shula, R. G. (2005). Basal area and diameter distribution prediction equations for young Douglas-Fir plantations with hardwood competition: Coast ranges. *Western Journal of Applied Forestry*, *20*(2), 77-93.

Laar, A. V., & Akça, A. (2007). *Forest mensuration*: in Managing Forest Ecosystems, Dordrecht, The Netherlands: Springer. 383 p.

Lee, Y. J., & Coble, D. W. (2006). A new diameter distribution model for unmanaged loblolly pine plantations in East Texas. *Southern Journal of Applied Forestry*, *30*(1), 13-20.

Lei, Y. (2008). Evaluation of three methods for estimating the Weibull distribution parameters of Chinese pine (*Pinus tabulaeformis*). *Journal of Forest Science*, *54*(12), 566-571.

Liu, C., Zhang, S. Y., Lei, Y., Newton, P. F., & Zhang, L. (2004). Evaluation of three methods for predicting diameter distributions of black spruce (*Picea mariana*) plantations in central Canada. *Canadian Journal of Forest Research*, *34*(12), 2424-2432.

Liu, C., Beaulieu J., Pregent G. & Zhang, S. Y. (2009). Applications and comparison of six methods for predicting parameters of the Weibull function in unthinned *Picea glauca* plantations. *Scandinavian Journal of Forest Research*, 24(1), 67–75.

Lohrey, R. E., & Bailey, R. L. (1977). *Yield tables and stand structure for unthinned longleaf pine plantations in Louisiana and Texas* (Vol. 133).

Department of Agriculture, USDA Forest Service, Southern Forest Experiment Station. Research Paper SO-133, 55 p.

Maltamo, M. (1997). Comparing basal area diameter distributions estimated by tree species and for the entire growing stock in a mixed stand. *Silva fennica*, 31, 53-65.

Matney, T. G., & Sullivan, A. D. (1982). Compatible stand and stock tables for thinned and unthinned loblolly pine stands. *Forest Science*, *28*(1), 161-171.

McTague, J. P., & Bailey, R. L. (1987). Compatible basal area and diameter distribution models for thinned loblolly pine plantations in Santa Catarina, Brazil. *Forest science*, *33*(1), 43-51.

Nelder, J.A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7, 308–313.

Nelson, T. C. (1964). Diameter distribution and growth of *Loblolly pine*. *Forest Science*, 10 (1), 105-114.

Ogana, F. N., Gorgoso-Varela, J. J., & Osho, J. S. A. (2020). Modelling joint distribution of tree diameter and height using Frank and Plackett copulas. *Journal of Forestry Research*, *31*(5), 1681-1690.

Packard K. C. (2000). Modeling tree diameter distributions for mixed-species conifer forests in the Northeast United States. Master Thesis. State Univ of New York College Environ Sci Forest. 129 pp.

Palahí, M., Pukkala, T., & Trasobares, A. (2006). Modelling the diameter distribution of *Pinus sylvestris*, *Pinus nigra* and *Pinus halepensis* forest stands in Catalonia using the truncated Weibull function. *Forestry*, *79*(5), 553-562.

Palahí, M., Pukkala, T., Blasco, E., & Trasobares, A. (2007). Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (north-east of Spain). *European Journal of Forest Research*, *126*(4), 563-571.

Podlaski, R. (2006). Suitability of the selected statistical distributions for fitting diameter data in distinguished development stages and phases of near-natural mixed forests in the Świętokrzyski National Park (Poland). *Forest Ecology and Management*, *236*(2-3), 393-402.

Poudel, K. P. (2011). Evaluation of methods to predict Weibull parameters for characterizing dimater distributions. Master Thesis, Graduate Faculty of the Louisiana State University and Agricultural and Mechanical Collage, 60 p., USA.

Poudel K.P., & Cao Q. V. (2013). Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. *Forest Sci*, 59(2), 243-252.

Pukkala, T., Saramaki, J., & Mubita, O., (1990). Management Planning System for Tree Plantations; A Case Study for *Pinus Kesiya* in Zambia. Silva Fennica, 24, 171–180.

R Core Team, (2024). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

Rennolls, K., Geary, D. N., & Rollinson, T. J. D. (1985). Characterizing diameter distributions by the use of the Weibull distribution, *Forestry*, 58, 58-66.

Saramäki, J., (1992) A growth and yield prediction model of *Pinus kesiya* (Royle ex Gordon) in Zambia. *Acta Forestalia Fenica*, 230, 68 p.

Scolforo, J. R. S., Tabai, F. C. V., de Macedo, R. L. G., Acerbi Jr, F. W., & de Assis, A. L. (2003). SB distribution's accuracy to represent the diameter distribution of Pinus taeda, through five fitting methods. *Forest ecology and management*, *175*(1-3), 489-496.

Smalley, G. W. & Bailey, R. L. (1974). Yield tables and stand structure for shortleaf pine plantations in the Tennessee, Alabama and Georgia Highlands. USDA Forest Service Research Paper, 97 p.

Weibull, W. (1951). A statistical distribution of wide applicability. *J. Appl. Mech.*, Vol. 18, p. 293-297.

Yavuz, H., Gül, A. U., Mısır, N., Özçelik, R. & Sakıcı, O. E. (2002). Meşcerelerde çap dağılımının düzenlenmesi ve bu dağılımlara ilişkin parametreler ile çeşitli meşcere öğeleri arasındaki ilişkilerin belirlenmesi. Orman Amenajmanı'nda Yeni Kavramsal Açılımlar ve Yeni Hedefler Sempozyumu, 2012, İstanbul.

Zhang, L., & Liu, C. (2006). Fitting irregular diameter distributions of forest stands by Weibull, modified Weibull, and mixture Weibull models. *Journal of Forest Research*, *11*(5), 369-372.

Chapter 8

Use of the Weibull Distribution in Modeling Diameter Distributions in Forestry

Muammer ŞENYURT[1]

Introduction

Sustainable management of natural resources is crucial for maintaining environmental balance and maximizing economic benefits. Forest ecosystems play a critical role in this context, and their accurate assessment requires the application of various analytical techniques. The distribution of tree diameters serves as a fundamental metric for understanding forest structure and functions. Diameter distributions provide insights into tree age, size, and diversity, and this information can be utilized in multiple aspects of forest management.

In forestry planning and related activities, obtaining accurate and detailed estimates of forest stands is of utmost importance. Diameter distribution models, together with yield tables, provide detailed information at the diameter step or diameter class level, enabling a better understanding of stand structure. These models make significant contributions to multi-purpose planning processes.

Ensuring the sustainability and operational efficiency of forest resource management necessitates a rigorous scientific assessment of growth dynamics and productivity capacity (Akalp 1983). Within the forestry discipline, modeling frameworks for growth and yield—constructed through datasets derived from both permanent and temporary sampling installations—serve as fundamental tools for evaluating the production capabilities of forested ecosystems (Vanclay 1994, Mısır 2003). Such modeling approaches are indispensable for projecting future yield trajectories, assessing the impacts of silvicultural interventions, and formulating comprehensive forest management strategies (Peng 2000). Growth modeling frameworks are conventionally classified into two primary categories: empirically-based and mechanistically-

[1]Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Kastamonu, Türkiye, Orcid: 0000-0002-8957-9295, msenyurt@karatekin.edu.tr

driven models. The empirical category is further differentiated into three subcategories according to the spatial and organizational scale of the modeling unit: stand-scale models, size-class distribution models, and single-tree models (Mısır 2003).

Within the framework of Forest Management Plans governing Turkey's forested landscapes, modeling systems that quantify stand increment and productivity constitute an essential foundation for planning processes. For even-aged forest systems, yield tabulations (referred to as Whole-Stand Models) function as instruments for deriving per-hectare growth and productivity metrics, while also establishing optimal structural configurations for stands. Put differently, these tabulations deliver aggregate stand volume estimates corresponding to specific age intervals. Nevertheless, such approaches exhibit limitations in addressing targeted inquiries, including the determination of cumulative volume for trees exceeding 19 cm in thickness or those falling below 40 cm in diameter within a given stand unit.

Models characterizing diameter distributions represent a critical element within the broader framework of size-class modeling approaches, facilitating the quantification of stem frequency across distinct diameter categories. Data resolution at the size-class scale proves particularly valuable for the strategic management of multi-aged forest systems. Through the application of diameter distribution modeling techniques, stem density within individual size classes can be quantified, while size-class transition models enable the projection of temporal shifts as trees progress into subsequent diameter categories. Consequently, such modeling frameworks hold substantial significance, particularly in defining optimal structural attributes for multi-aged forest ecosystems and in identifying specific stem populations requiring silvicultural treatment under prevailing stand conditions.

Within forestry science, size-class modeling approaches serve as a bridging mechanism between aggregate stand-scale models and single-tree modeling frameworks, operating through the characterization of stem frequency patterns across diameter categories within forest stands (Vanclay 1994). Given that stem diameter exerts direct influence on harvestable volume output and product assortment diversity (Kalıpsız 1998), models depicting diameter distributions assume critical strategic value in forest resource management and planning processes. The analysis of diameter distributions enhances comprehension of structural attributes within stands by elucidating productivity patterns and developmental trajectories of tree cohorts spanning various size classes. Such analytical insights prove especially pertinent when evaluating prospective

influences of silvicultural interventions on stand developmental processes (Gorgoso et al. 2012).

Within the construction framework of diameter distribution modeling systems, probability density functions serve as the predominant analytical tool. Several statistical distributions, including Weibull, Gamma, Log-normal, Beta, and Johnson's SB, have gained particular prominence in forestry applications (Liu et al. 2009). The Weibull function was initially introduced to forest science through the pioneering work of Bailey and Dell (1973) and has subsequently emerged as a fundamental tool for characterizing diameter distribution patterns with considerable flexibility (Merganič and Sterba 2006; Stankova and Zlatanov 2010; Diamantopoulou et al. 2015). The extensive adoption of the Weibull function stems from multiple advantageous attributes: its demonstrated efficacy across both uniform-aged and multi-aged forest structures, the computational simplicity associated with parameter derivation, and its remarkable versatility in capturing diverse distributional forms (Bolat 2014).

Within forest science contexts, multiple parameter estimation techniques have emerged for the Weibull function, encompassing maximum likelihood estimation, method of moments, quantile-based approaches, and combined methodologies. Quantile-based estimation techniques have gained increasing adoption in contemporary research, attributed to their perceived capacity to capture fundamental ecological dynamics (Borders et al. 1987; Liu et al. 2004).

In Turkey, yield tables prepared using traditional methods have been widely used for many years. However, since these tables provide only generalized data, diameter distribution models offer significant advantages for forestry planning by enabling more detailed analyses (Bolat 2015). Estimating stand characteristics based on diameter classes is important not only for accurately determining the economic value of forest products but also for conducting detailed analyses tailored to different management and utilization scenarios (Packard 2000).

In recent years, the implementation of forest planning within the framework of an ecosystem-based functional planning approach has further increased the importance of diameter distribution models (Huang et al. 2000; Bankston et al. 2021). Such models provide reliable tools for predicting both the current and future states of forest resources, thereby playing a critical role in achieving sustainable forest management objectives (Bailey and Dell 1973; Vanclay 1994). Moreover, the detailed characterization of diameter distribution models across different diameter ranges enables more precise planning of silvicultural interventions and supports accurate evaluation of ecological parameters such as forest ecosystem carbon stocks (Chen 2004; Özçelik et al. 2016).

For forests to be managed and planned to serve multiple purposes, it is essential to accurately estimate the growth and yield values, as well as various stand characteristics, across stands with differing attributes. In this context, the estimation of stand growth components and related values holds great importance for diverse forestry practices and planning activities.

To ensure optimal planning of Turkey's forests, the development of appropriate decision support systems requires detailed information about forest stands and the diameter classes that constitute these stands. Therefore, through this study, estimates of growth, increment, and various growth components obtained at the diameter-class level will provide more detailed information compared to traditional yield tables, offering insights at the diameter step or diameter class scale. Furthermore, the study will examine how the parameters of diameter distributions vary with changes in stand mixture ratios and developmental stages. The results obtained are expected to complement existing scientific research in this field and contribute meaningfully to the related literature.

Successful implementation of diameter distribution modeling frameworks hinges not solely upon judicious model selection, but equally upon the rigorous and unbiased derivation of model coefficients (Poudel and Cao 2013). To achieve this objective, coefficient estimation approaches are conventionally organized into two primary groupings: techniques employing empirical associations that facilitate immediate coefficient derivation, and approaches that extract coefficients through diameter statistical moments or quantile values (Hyink and Moser 1983; Siipilehto et al. 2007).

Research endeavors within Turkey have demonstrated the viable implementation of diameter distribution modeling across diverse forest stand configurations and tree taxa, with statistical functions including Weibull, Johnson SB, Gamma, and Log-normal proving capable of capturing distributional patterns effectively (Carus 1996; Ercanlı and Yavuz 2010; Sakıcı and Gülsunar 2012; Bolat and Ercanlı 2017; Sakıcı and Dal 2021). Within the scientific literature, a consensus has emerged that stand-specific attributes serve as primary determinants governing the optimal selection among alternative distributional functions (Liu et al. 2014; Wang and Rennolls 2005).

Within forest science, diameter distribution modeling constitutes a critical analytical tool essential for achieving sustainable resource stewardship, conducting robust economic evaluations, and formulating silvicultural intervention strategies. Future investigative efforts should prioritize the advancement of refined estimation techniques for diameter distributions while

simultaneously improving methodological adaptability across heterogeneous stand conditions.

Investigative work concerning diameter distribution characterization commenced during the closing decades of the 1800s, with progressive refinements in statistical theory subsequently enabling the creation of increasingly complex modeling frameworks. The pioneering contributions of Gram (1883) and De Liocourt (1898) established the initial diameter distribution models utilizing normal and exponential functions, respectively (Leak 1965). By the 1960s, probability density functions had achieved widespread adoption in forestry applications, with the Beta function being introduced to the discipline through the seminal work of Clutter and Bennet (1965) (Packard 2000).

Turkish research initiatives in this domain have predominantly concentrated on species-specific analyses. Illustratively, Saraçoğlu (1988) implemented Meyer's exponential formulation for multi-aged fir populations, whereas Carus (1996) documented the efficacy of the Gamma function in uniform-aged beech ecosystems. Contemporary investigations have revealed superior performance characteristics for Johnson's SB and Weibull functions across heterogeneous stand configurations (Ercanlı and Yavuz 2010; Kahriman and Yavuz 2011).

Probability Density Functions

Probability Density Functions (PDFs), constituting a fundamental component within statistical theory, find application in forest science for characterizing stem frequency across size categories and constructing diameter distribution modeling frameworks. Such functions quantify the proportional representation of stems within designated diameter intervals relative to the aggregate stem population, yielding probability estimates spanning the zero-to-unity interval. Multiple probability density functions have demonstrated utility for these applications, including the Normal (Bailey 1980), Lognormal (Bliss and Reinker 1964), Gamma (Nelson 1964), Beta (Clutter and Bennet 1965; Zöhrer 1969), Johnson's SB (Johnson 1949), and Weibull formulations (Weibull 1951; Bailey and Dell 1973).

Throughout the evolution of diameter distribution characterization, diverse probability density functions have undergone empirical evaluation. Nevertheless, accumulated evidence demonstrates that the three-parameter Weibull probability density function yields robust outcomes when applied to diameter distribution modeling (Smalley and Bailey 1974; Lohrey and Bailey 1977; Feduccia et al. 1979; Matney and Sullivan 1982; Clutter et al. 1984; Baldwin and Feduccia 1987; Liu et al. 2004; Čavlović et al. 2006; Palahi et al.

2006; Zhang and Liu 2006; Gorgoso-Varela et al. 2007; Jiang and Brooks 2009; Andrasev et al. 2009). Consequently, the three-parameter Weibull formulation has emerged as the preferred analytical tool for characterizing diameter distributions within forest stands. While alternative probability density functions have historically been utilized and remain in contemporary practice, the three-parameter Weibull function has garnered increasing prominence in recent decades, demonstrating enhanced performance attributed to its exceptional capacity for representing ecological heterogeneity. Empirical investigations conducted by Bailey and Dell (1973), Rennols et al. (1985), Knoebel et al. (1986), Pukkala et al. (1990), Cao (2004), Palahi et al. (2006), and Nord-Larsen et al. (2006) exemplify research endeavors wherein probability density functions served as analytical instruments for diameter distribution determination. The mathematical specification of the three-parameter Weibull probability density function appears as follows (Bailey and Dell 1973):

$$F(x,\alpha,\beta,\gamma) = \frac{\alpha}{\beta}\cdot\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\cdot exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)$$

Various coefficient estimation formulations serve to derive the coefficients (γ, α, β) characterizing probability density functions. Within the Weibull formulation, the location coefficient (α) designates the minimum diameter threshold within the distribution, the scale coefficient (β) governs the distributional dispersion, and the shape coefficient (γ) dictates the distributional morphology. Figure 1 demonstrates the behavioral response of the Weibull probability density function to modifications in the scale coefficient (β), whereas Figure 2 presents alterations associated with the shape coefficient (γ).

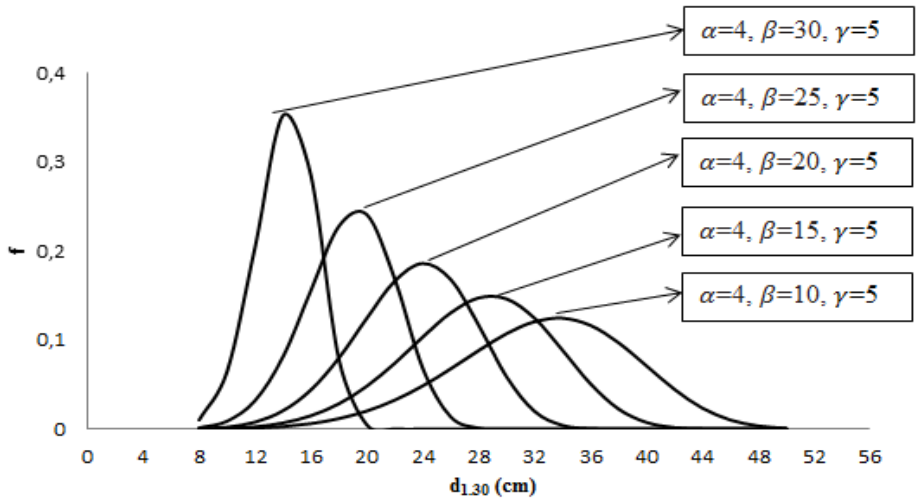

Figure 1. Variation of the Weibull probability density function with respect to the scale parameter (β) (Bolat 2014)

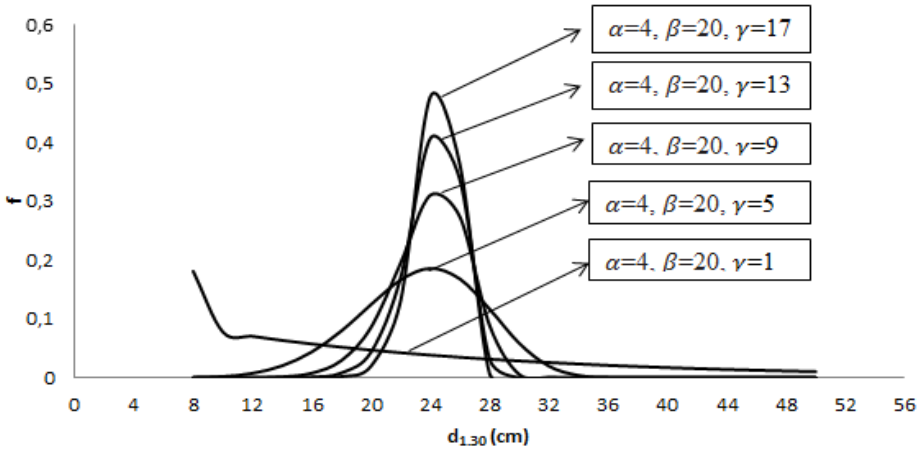

Figure 2. Variation of the Weibull probability density function with respect to the shape parameter (γ) (Bolat 2014)

Parameter Estimation Methods

For deriving the coefficients of the three-parameter Weibull probability density function, multiple statistical methodologies are available for implementation. In this regard, techniques including Nonlinear Regression Analysis, Maximum Likelihood Estimation, Method of Moments, and Quantile-Based Coefficient Recovery have achieved extensive adoption within the scientific literature (Bolat, 2014).

Within the array of techniques employed for deriving coefficients of probability density functions, formulations grounded in statistical moments and quantile values of diameter distributions occupy a prominent position. Within this framework, investigations by Cao (2004) and Poudel and Cao (2013) represent notable contributions to research examining the three-parameter Weibull formulation.In these referenced investigations, moment-based formulations incorporated the minimum threshold value, variance, arithmetic mean diameter ($\overline{D}$), and the quadratic mean diameter, which corresponds to the diameter of the tree representing the median basal area within the diameter distribution. Conversely, quantile-based formulations utilized diameter values associated with the 31st, 50th, 63rd, and 95th quantiles, derived through ascending rank-ordering of diameter observations.

Additionally, within a combined methodology designated as the hybrid technique, both moment-derived and quantile-derived values were integrated concurrently. This technique employed a modeling framework that incorporates the arithmetic mean diameter, the quadratic mean diameter, alongside diameter values associated with the 25th, 31st, 50th, 63rd, and 95th quantiles.

In this context, eight different methods can be applied to estimate the parameters of the three-parameter Weibull function:

Method (1): The Maximum Likelihood Estimation (MLE) based on the Nelder–Mead optimization algorithm,

Method (2): An equation utilizing the variance and arithmetic mean diameter ($\overline{D}$) values derived from the moments of the diameter distribution,

Method (3): A moment-based equation employing the arithmetic mean diameter ($\overline{D}$) and the quadratic mean diameter (D_g),

Method (4): A percentile-based equation using diameter values corresponding to the 31st and 63rd percentiles,

Method (5): Another percentile-based equation employing the diameters corresponding to the 50th and 95th percentiles,

Method (6): A hybrid equation that combines the quadratic mean diameter (D_g) with the diameters corresponding to the 25th and 95th percentiles (Poudel and Cao, 2013),

Method (7): A hybrid equation utilizing the quadratic mean diameter (D_g) together with the 31st and 63rd percentile diameters,

Method (8): A hybrid equation that simultaneously incorporates the quadratic mean diameter (D_g) and the diameters corresponding to the 25th, 50th, and 95th percentiles.

The formulas of these methods related to the three-parameter Weibull function, as proposed by Cao (2004) and Poudel and Cao (2013), are presented in below Table

Table Equations for estimating the parameters of the three-parameter Weibull function using moment-based and percentile-based methods (Cao, 2004; Poudel and Cao, 2013)

Parameter Estimation Method	**Equalities**
Method (1): Maximum Likelihood Method	$\frac{\sum_{i=1}^{n}(x_i)^{\gamma} ln(x_i)}{\sum_{i=1}^{n}(x_i)^{\gamma}} - \frac{1}{\gamma} = \frac{1}{n}\sum_{i=1}^{n}(x_i)$ $\beta = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i)^{\gamma}\right)^{\frac{1}{c}}$
Method (2): Moment method including the variance value and arithmetic mean ($\overline{D}$) values of the diameters	$\alpha = 0.5 \cdot d_{min}$ $\beta = \frac{\overline{D}-\alpha}{G_1}$ γ parameter is obtained from the solution $\beta^2(G_2 - G_1^2) - d_{var} = 0$
Method (3): Moment method including arithmetic mean ($\overline{D}$) and median diameter (D_g) values	$\alpha = 0.5 \cdot d_{min}$ $\beta = -\alpha \cdot \left(\frac{G_1}{G_2}\right) + \left[\left(\frac{\alpha}{G_2}\right)^2 \cdot (G_1^2 - G_2) + \frac{D_g^2}{G_2}\right]^{0.5}$ γ parameter is obtained from the solution $\beta^2(G_2 - G_1^2) - d_{var} = 0$
Method (4): Percentile method including diameter values corresponding to the 31% and 63% percentile values	$\alpha = 0.5 \cdot d_{min}$ $\gamma = \frac{Ln\left(\frac{Ln(1-0.63)}{Ln(1-0,31)}\right)}{Ln(d_{\%63}-\alpha)-Ln(d_{\%31}-\alpha)}$ $\beta = \frac{d_{\%63} - \alpha}{\left(-Ln(1-0.63)\right)^{\frac{1}{\gamma}}}$
Method (5): Percentile method including diameter values corresponding to the 50% and 95% percentile values	$\alpha = 0.5 \cdot d_{min}$ $\gamma = \frac{Ln\left(\frac{Ln(1-0.95)}{Ln(1-0,50)}\right)}{Ln(d_{\%95}-\alpha)-Ln(d_{\%50}-\alpha)}$ $\beta = \frac{d_{\%50} - \alpha}{\left(-Ln(1-0.50)\right)^{\frac{1}{\gamma}}}$
Method (6): Hybrid method including diameter values corresponding to the median diameter (D_g), 25% and 95%	$\alpha = 0.5 \cdot d_{min}$ $\gamma = \frac{Ln\left(\frac{Ln(1-0.95)}{Ln(1-0,25)}\right)}{Ln(d_{\%95}-\alpha)-Ln(d_{\%25}-\alpha)}$

percentile values	$\beta = -\alpha \cdot \left(\frac{G_1}{G_2}\right) + \left[\left(\frac{\alpha}{G_2}\right)^2 \cdot (G_1^2 - G_2) + \frac{D_g^2}{G_2}\right]^{0.5}$
Method (7): Hybrid method with diameter values corresponding to the 31% and 63% percentile values of the median diameter (D_g)	$\alpha = 0.5 \cdot \mathrm{d}_{min} \quad \gamma = \frac{Ln\left(\frac{Ln(1-0.63)}{Ln(1-0,31)}\right)}{Ln(d_{\%63}-\alpha)-Ln(d_{\%31}-\alpha)}$ $\beta = -\alpha \cdot \left(\frac{G_1}{G_2}\right) + \left[\left(\frac{\alpha}{G_2}\right)^2 \cdot (G_1^2 - G_2) + \frac{D_g^2}{G_2}\right]^{0.5}$
Method (8): Hybrid method including diameter values corresponding to the 25%, 50% and 95% percentile values of the median diameter (D_g)	$\alpha = \frac{n^{0.3333} \cdot d_{min} - d_{\%50}}{n^{0.3333} - 1}$ $\gamma = \frac{Ln\left(\frac{Ln(1-0.95)}{Ln(1-0,50)}\right)}{Ln(d_{\%95}-\alpha)-Ln(d_{\%50}-\alpha)}$ $\beta = -\alpha \cdot \left(\frac{G_1}{G_2}\right) + \left[\left(\frac{\alpha}{G_2}\right)^2 \cdot (G_1^2 - G_2) + \frac{D_g^2}{G_2}\right]^{0.5}$

In the equations; n; the number of trees in the sample area, ; the minimum diameter in the sample area, d_{min}; the minimum diameter, d_{var}; the variance of the diameter values, $d_{\%25}$, $d_{\%31}$, $d_{\%50}$ $d_{\%63}$, $d_{\%95}$; the diameter value corresponding to the 25%, 31%, 50%, 63% and 95% data when the data are sorted from smallest to largest in the diameter distribution, d_g; the diameter of the basal area middle tree calculated for the sample area, and G_1 and G_2; show the value corresponding to the Gamma function.

Lohrey and Bailey (1977) employed diameter values associated with the 24th and 93rd quantiles for deriving Weibull function coefficients in diameter distribution characterization. McTague and Bailey (1987) incorporated diameters at the 10th, 63rd, and 93rd quantiles for coefficient derivation. Bailey (1989) conducted coefficient estimation for the Weibull function utilizing the minimum diameter threshold, the quadratic mean diameter (dg), alongside diameter values at the 25th, 50th, and 95th quantiles. Bullock and Burkhart (2005) derived coefficients through the quadratic mean diameter combined with 25th and 97th quantile diameter values. In a comparative investigation examining diameter distribution modeling performance across alternative quantile selections, Poudel (2011) demonstrated that the approach integrating the quadratic mean diameter (dg) with 25th, 50th, and 95th quantile diameters yielded optimal outcomes.

Correspondingly, multiple investigations employing varied quantile intervals for deriving Weibull function coefficients (Brooks et al. 1992; Knowe et al. 2005; Lee and Coble 2006; Coble and Lee 2008; Jiang and Brooks 2009) have achieved effective diameter distribution characterization. Variables including

stand age, site productivity, stocking density, and micro-environmental attributes may induce substantial structural heterogeneity within diameter distributions. Consequently, the optimal selection of quantile intervals for accurately capturing diameter distribution patterns, along with the efficacy of coefficient derivation techniques grounded in these quantiles, exhibits context-dependency relative to site-specific conditions and stand compositional attributes. A quantile configuration demonstrating robust performance at one location may exhibit inadequate representational capacity at alternative sites.

Sakıcı and Dal (2021) established that, within a comparative evaluation of alternative probability density functions for characterizing diameter distributions across monospecific Scots pine (Pinus sylvestris L.) forests in the Kastamonu region, the Johnson SB function exhibited superior performance, attaining a mean success ranking of 1.75. Moreover, upon investigating performance variability in relation to stand attributes, the analysis revealed that stem frequency, average diameter, crown cover percentage, successional stage, and stand classification exerted influence on functional rankings, while basal area density and site index class demonstrated negligible impact on ranking outcomes.

Diameter distribution models are of great importance for providing more detailed predictions of stand structures and for determining the types of products obtainable from forests. In Turkey, growth and yield estimates are currently based on conventional yield tables and frequency-based yield tables. However, these tables generally provide only generalized estimates for the stand. In contrast, many forestry activities, particularly forest management planning, require more detailed analyses of stand structures.

References

Akalp, T. 1983. Değişik Yaşlı Meşcerelerde Artım ve Büyümenin Simülasyonu, İ.Ü. Yayın No:3051, O.F. Yayın no: 327, İstanbul, 169 s.

Andrasev, S., Bobinac, M. and Orlovic, S. 2009. Diameter structure models of Black Poplar selected clones in the section Aigeiros (Duby) obtained by the Weibull distribution. Sumarski List, 133: 589-603

Bailey, R. L. and Dell, T.R. 1973. Quantifying diameter distributions with the weibull function. Forest Science (19): 97-104.

Bailey, R. L. 1980. Individual tree growth derived from diameter distribution models. Forest Science, 26(4): 626-632.

Bailey, R.L., Burgan, T.M. and Jokela, E.J. 1989. Fertilized midrotation-aged Slash pine plantations-stand structure and yield prediction models. Southern Journal of Applied Forestry, 13(2): 76-80.

Baldwin, V. C. and Feduccia, D. P. 1987. Loblolly pine growth and yield prediction for managed West Gulf plantations. USDA For. Serv. Res. Pap. SO 236, 27 p.

Bankston, J.B., Sabatia, C.O. and Poudel, K.P. 2021. Effects of sample plot size and prediction models on diameter distribution recovery. Forest Science, 67(3): 245-255.

Bliss, C.I. and Reinker, K.A. 1964. A lognormal approach to diameter distributions in even-aged stands. Forest Science, 10(3): 350-360.

Bolat F. 2014. Bursa-Kestel Orman İşletme Şefliği içerisindeki meşcereler için çap dağılım modellerinin geliştirilmesi. Yüksek Lisans Tezi, Çankırı Karatekin Üniversitesi, 77 sayfa, Çankırı.

Bolat, F. 2015. Predictions for Oriental Beech Tree Heights Based on Artificial Neural Network in Kestel Forests, The 10th International Beech Symposium, 1-6 September, 2015, Kastamonu, Turkey.

Bolat F. and Ercanlı İ. 2017. Modeling diameter distribution by using Weibull function in forests located Kestel-Bursa. Kastamonu University Journal of Forestry Faculty, 17(1), 107-115.

Borders, B.E., Souter, R.A., Bailey, R.L. and Ware, K.D. 1987. Percentile based distributions characterize forest tables. Forest Science, 33(2): 570-576.

Brooks, J. R., Borders, B. E. and Bailey, R. L. 1992. Predicting diameter distributions for site prepared loblolly and slash pine plantations. Southern Journal of Applied Forestry, 16: 130-133.

Bullock, B. P., and Burkhart, H. E. 2005. Juvenile diameter distributions of loblolly pine characterized by the two-parameter Weibull function. New Forests, 29: 233-244.

Cao, Q.V. 2004. Predicting parameters of a weibull function for modelling diameter distribution. Forest Science (50): 682-685.

Carus. S. 1996. Aynı Yaşlı Doğu Kayını (Fagus orientalis Lipsly.) Meşcerelerinde çap dağılımın bonitet ve yaşa göre değişimi. İstanbul Orman Fakültesi Dergisi, 46: 171-181.

Čavlović, J., Božić, M. and Boncina, A. 2006. Stand Structure of An Uneven-Aged Fir–Beech Forest with An Irregular Diameter Structure: Modeling The Development of The Belevine Forest, Croatia. European Journal of Forest Research, 125 (4): 325-333.

Chen, X. 2004. Using Weibull distribution in diameter distributions modeling. Journal of Forestry Research, 15(2): 150-154.

Clutter, J.L. and Bennet, F.A. 1965. Diameter distributions in old-field slash pine plantation. Georgia Forest Research Council, Report No: 13, USA.

Clutter, J.L., Harms, W.R., Brister, G.H. and Rhenney, J.W. 1984. Stand structure and yields of site-prepared loblolly pine plantations in the lower coastal plain of the Carolinas, Georgia, and North Florida. USDA For. Serv. Gen. Tech. Rep. SE-27, 173 s.

Coble, D. W. and Lee, Y. J. 2008. A new diameter distribution model for unmanaged slash pine plantations in East Texas. Southern Journal of Applied Forestry, 32(2): 89-94.

Diamantopoulou, M. J., Özçelik, R., and Crecente-Campo, F. 2015. Weibull parameters estimation using a hybrid method. Forest Systems, 24(1), e001

Ercanlı, İ. ve Yavuz, H. 2010. Doğu ladini (Picea Orientalis (L.) Link)-Sarıçam (Pinus Sylvestris L.) karışık meşcerelerinde çap dağılımlarının olasılık yoğunluk fonksiyonları ile belirlenmesi. Kastamonu Orman Fakültesi Dergisi, 10(1): 68-83.

Feduccia, D. P., Dell, T. R., Mann, W. F. and Polmer, B. H. 1979. Yields of unthinned loblolly pine plantations on cutover sites in the West Gulf region. USDA For. Serv. Res. Pap. So-148, 88 p.

Gorgoso-Varela, J. J., Alvarez-Gonzalez, J. G., Rojo, A. and Grandas-Arias, J. A. 2007. Modelling diameter distributions of Betula alba L. stands in northwest Spain with the two-parameter Weibull function. Invest. Agrar. Sist. Recur. For. 16: 113-123.

Gorgoso, J.J. Rojo, A., Camara-Obregon, A. and Dieguez-Aranda, U. 2012. A Comparison of Estimation Methods for Fitting Weibull, Johnson's SB and Beta Functions to Pinus pinaster, Pinus radiata and Pinus sylvestris Stands in Northwest Spain. Forest Systems 21(3): 446-459.

Huang, S., Price, D., Morgan, D. and Peck, K., 2000. Kozak's variable exponent taper equation regionalized for white spruce in Alberta. Western Journal of Applied Forest, 15: 75-85.

Hyink, D.M. and Moser, J.W. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. Forest Science, 29(1): 85-95.

Jiang, L. C. and Brooks, J. R. 2009. Predicting diameter distributions for young longleaf pine plantations in Southwest Georgia. Southern Journal of Applied Forestry, 33: 25-28.

Johnson, N.L. 1949. System of frequency curves generated by methods of translation. Biometrika (36); 149-176.

Kahriman, A. ve Yavuz, H. 2011. Sarıçam (Pinus sylvestris L.)-Doğu kayını (Fagus orientalis Lipsky) karışık meşcerelerinde çap dağılımlarının olasılık yoğunluk fonksiyonları ile belirlenmesi. Artvin Çoruh Üniversitesi Orman Fakültesi Dergisi, 12 (2):109-125.

Kalıpsız, A. 1998. Orman hasılat bilgisi. İstanbul Üniversitesi Yayınları, İstanbul.

Knoebel, B.R., Burkhart, H.E. and Beck, D.E. 1986. A growth and yield model for thinned stands of yellow-poplar. Forest Science Monograph (27): 39 p.

Knowe, S.A. 1992. Basal area and diameter distribution models for loblolly pine plantations with hardwood competition in the piedmont and upper coastal plain. Southern Journal of Applied Forestry (16): 93-98.

Knowe, S. A., Ahrens, G. A., and DeBell, D. S. 1997. Comparison of diameter-distribution prediction, stand table projection and individual-tree growth modeling approaches for young red alder plantations. Forest Ecology and Management, 98(1): 49-60.

Knowe, S. A., Radosevich, S. R., and Shula, R. G. 2005. Basal area and diameter distribution prediction equations for young Douglas-Fir plantations with hardwood competition: Coast ranges. Western Journal of Applied Forestry, 20(2), 77-93.Leak, W. B. 1965. The J-Shaped Probability Distribution. Forest Science, 11: 405–409.

Leak, W. B. (1965). The J-shaped probability distribution. Forest Science, 11(4), 405-409.

Lee, Y. J., and Coble, D. W. 2006. A new diameter distribution model for unmanaged loblolly pine plantations in East Texas. Southern Journal of Applied Forestry, 30(1): 13-20.

Liu, C., Zhang, S.Y., Lei, Y., Newton, P.F. and Zhang, L. 2004. Evaluation of tree methods for predicting diameter distributions of black spruce (Picea mariana) plantations in central Canada. Canadian Journal of Forest Research, 34(12): 2424-2432.

Liu, C., Beaulieu J., Pregent G. and Zhang, S.Y. 2009. Applications and comparison of six methods for predicting parameters of the Weibull function in unthinned Picea glauca plantations. Scandinavian Journal of Forest

Research, 24(1): 67–75.
Liu, F., Li, F., Zhang, L., Jin, X., 2014. Modeling diameter distributions of mixed-species forest stands. Scandinavian Journal of Forest Research, 29(7): 653-663.
Lohrey, R.E. and Bailey, R. L. 1977. Yield tables and stand structure for unthinned long leaf pine plantations in Louisiana and Texas. USDA For. Ser. Res. Pap. SO-133. 55 s
Matney, T. G. and Sullivan, A. D. 1982. Compatible Stand and Stock Tables for Thinned and Unthinned Loblolly Pine Stands. Forest Science, 28: 161–171.
McTague, J. P. and Bailey, R. L. 1987. Compatible basal area and diameter distribution models for thinned loblolly pine plantations in Santa Catarina, Brazil. Forest science, 33(1), 43-51.
Merganič, J. and Sterba, H. 2006. Characterisation of diameter distribution using the Weibull function: method of moments. European Journal of Forest Research, 125: 427-439.
Mısır, N. 2003. Karaçam ağaçlandırmalarına ilişkin büyüme modelleri. Doktora tezi, KTÜ Fen Bilimleri Enstitüsü, 222 s., Trabzon.
Nelson, T.C. 1964. Diameter distribution and growth of Loblolly pine. Forest Science, 10(1); 105-114.
Nord-Larsen, T. and Cao, Q.V. 2006. A diameter distribution model for even-aged beech in Denmark. Forest Ecology and Management (231): 218-225.
Özçelik, R., Fidalgo Fonseca, T. J., Parresol, B. R., and Eler, Ü. 2016. Modeling the diameter distributions of Brutian pine stands using Johnson's SB distribution. Forest Science, 62(6), 587-593.
Packard, K. C. 2000. Modeling Tree Diameter Distributions for Mixed-Species Conifer Forests in the Northeast United States. Master Thesis, State University of New York, New York, USA., 129 p.
Palahi, M., Pukkala, T., and Trasobares, A. 2006. Calibrating predicted tree diameter distributions in Catalonia (Spain). Silva Fennica (40-3): 487-500.
Peng, C., 2000. Growth and Yield Models for Uneven-Aged Stands: Past, Present and Future. Forest Ecology and Management 132(2-3): 259-279.
Poudell, K.P. 2011. Evaluation of methods to predict Weibull parameters for characterizing dimater distributions. Master Thesis, Graduate Faculty of the Louisiana State University and Agricultural and Mechanical Collage, 60 p., USA.
Poudell, K.P. and Cao Q.V. 2013. Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. Forest Science, 59(2); 243-252.

Pukkala, T., Saramaki, J., and Mubita, O. 1990. Management planning system for tree plantations; A case study for Pinus kesiya in Zambia. Silva Fennica (24); 171-180.

Rennolls K. Geary, D.N. and Rollinson, T.J.D. 1985. Characterizing diameter distributions by the use of the weibull distributions. Forestry (58): 58-66.

Sakıcı, O. E. and Gülsunar, M. 2012. Diameter distribution of Bornmullerian fir in mixed stands. Kastamonu Üniversitesi, Orman Fakültesi Dergisi, 12(3): 263-270.

Sakıcı, O. E. ve Dal, E. 2021. Kastamonu yöresi sarıçam meşcereleri için çap dağılımlarının modellenmesi ve çeşitli meşcere özellikleri ile ilişkilerinin belirlenmesi. Bartın Orman Fakültesi Dergisi, 23(3): 1026-1041.

Saraçoğlu, Ö. 1988. Karadeniz yöresi Göknar meşcerelerinde artım ve büyüme. Orman Genel Müdürlüğü Yayınları, No: 25, 312 s., Ankara.

Smalley, G.W. and Bailey, R.L. 1974. Yield tables and stand structure for Shortleaf pine plantations in The Tennessee, Alabama and Georgia Highlands. Forest Service Research Paper, 97 p., USA.

Stankova, T. V., and Zlatanov, T. M. 2010. Modeling diameter distribution of Austrian black pine (Pinus nigra Arn.) plantations: a comparison of the Weibull frequency distribution function and percentile-based projection methods. European journal of forest research, 129: 1169-1179.

Siipilehto, J., Sarkkola, S., Mehtätalo, L., 2007. Comparing regression estimationtechiques when predicting diameter distribution of Scots pine on drainedpeatlands. Silva Fennica, 41(2): 333349.

Vanclay, J.K. 1994. Modelling forest growth: Applications to mixed tropical forests. CAB International, Department of Economics and Natural Resource, Royal Veterinary and Agricultural University, 312 p., Copenhagen, Denmark.

Wang, M., Rennolls, K. 2005. Tree diameter distribution modelling: Introducing the logit logistic distribution. Canadian Journal of Forest Research, 35(6): 1305-1313.

Weibull, W. 1951. A statistical distribution function of wide applicability. Journal of Applied Mechanics, (18): 293–297.

Zhang, L. and Liu, C. 2006. Fitting Irregular Diameter Distributions of Forest Stands By Weibull, Modified Weibull and Mixture Weibull Models. Journal Forest Research, 11: 369–372.

Zöhrer, F. 1969. Ausgleich von Häufigkeitsverteilungen mit hilfe der beta-funktion. Forstarchiv, 40(3): 37-42.

Chapter 9

Decision Science and Various Applications in Forestry

Bahadir Cagri BAYRAM[1]

Introduction

Decision-making is one of the most critical skills for addressing and solving complex, interwoven problems, which can take various forms, such as choices, rankings, acceptances, evaluations, and constructions (Jonassen, 2012). Nature-related issues have been ongoing since the existence of mankind on Earth. In the early days, we most likely struggled to decide how to utilize nature to ensure safe accommodation, locate the right water resources, and sustain ourselves while settling in the right place. With the evolution of mankind, our problems have also evolved. Simple lives have simple problems and simple solutions. However, today we have high-end industries and technologies that demand many sacrifices in every aspect of life. Due to the demanding needs of society and our modern world, we face some challenging and complex decisions. Hence, we need to employ robust methods to deal with these problems and provide the simplest solutions that we are capable of. Environmental issues are the ones we encounter most frequently in relation to nature. Forestry and related decisions are among the primary concerns that practitioners, strategists, and researchers typically face in environmental issues. To deal with such complex and multifaceted problems, we must employ modern numerical analysis approaches, such as decision support systems.

In this chapter, the decision-making process and its background, decision support systems, Multi-Criteria Decision Making (MCDM), and various applications in forestry and the forest products industry will be discussed. The main goal of this chapter is to prove that we can adopt these approaches and some particular methods to address our future “forestry problems” in different aspects.

[1] Kastamonu University, Faculty of Forestry, Forest Industry Engineering Department, Kastamonu, Türkiye, Orchid:0000-0002-8563-0233, bcbayram@kastamonu.edu.tr

Decision Science

Decisions involve a type of problem that is somewhat ill-structured, where the decision maker weighs two or more alternative outcomes and chooses one. Within a rational framework, the optimal choice is defined as the one that yields the highest utility. From a naturalistic orientation, the preferred option might be the one that best corresponds to an individual's convictions or historical data. Decisions, whether straightforward or complex, frequently involve numerous alternatives, diverse parameters for evaluation, and a multiplicity of viewpoints pertaining to each parameter (Jonassen, 2012). Decision-making is a staple for complex problems; however, it can also be applied to simple decisions. The primary difference between the two lies in how they will be applied. While some problems require only regular decision-making, more complex and interwoven problems demand sequential or iterative decision-making methods (Jonassen, 2000).

A core aspect of decision-making is the deliberate choice of one or more advantageous or acceptable options from a wide range of possibilities. Regardless of whether these options represent demands, methodologies, occurrences, forecasts, or prospects, every decision necessitates a steadfast commitment to an operational strategy designed to yield favorable outcomes for identified stakeholders (Yates, 2003). Choices, Acceptances/rejections, Evaluations, and Constructions are types of decisions. To be exact…

Choices: The process of selecting a smaller group from a larger collection of options (for example, choosing a smartphone to buy).

Acceptances/Rejections involve a dichotomous decision where only one particular option is either approved or denied (such as being admitted to a graduate program).

Evaluations: Assertions of value that are supported by promises to take action (how much you are prepared to offer for a land).

Constructions: Efforts to formulate optimal solutions based on current resources (for instance, determining a budget for financing a community organization) (Yates and Tschirhart, 2006).

The literature presents two primary perspectives on decision-making: normative (or prescriptive) models and descriptive (or naturalistic) models. Normative theories operate under the assumption that decision-makers are rational beings who strive to find the best choice that maximizes their utility in uncertain conditions. These theories establish guidelines or principles for making decisions. They focus on developing frameworks to help rational and informed individuals determine the best option, often by assigning numerical values to various choices. On the other hand, descriptive or naturalistic models examine

how decisions are actually made in practice. These models acknowledge that people are rarely as rational as normative theories predict. Instead, their decision-making is often influenced by unconscious biases, emotions, and past experiences (Jonassen, 2012).

Normative (Prescriptive) Models

Normative models can be categorized into three types: rational choice, cost-benefit analysis, and risk assessment. All these models aim to maximize the expected utility or value associated with a decision. Normative theories focus on identifying the optimal choice that maximizes utility in uncertain situations. These theories serve as the foundation of operations research and decision theory. Research in normative decision-making investigates how individuals should make choices involving risk or rationality within various theoretical frameworks. These studies have concentrated on developing theories about how rational, informed individuals select the best option, which is typically evaluated through statistical analysis of alternatives (Jonassen, 2012).

Rational Choice Models

As stated in its name, the rational choice methods are used to compare choices, and they are not helpful for evaluations and acceptances. It is also known as multi-attribute utility theory. The rational choice approach suggests evaluating various alternatives and logically choosing the one that has the highest points. There have been various rational choice models, including decision matrices, SWOT analysis, and force-field analysis (Armstrong, 1982; Lewin, 1943). Rational choice methods depend on quantitative analysis, in which weights are assigned and the weighted options are summed to ensure an objective, rigorous, and reliable decision can be made (Klein, 1998).

Cost-Benefit Analysis

It is a variant of rational choice models. Cost-benefit analysis is a systematic and analytical process used to compare the benefits and costs associated with evaluating decisions, particularly those that have social implications (Mishan and Quah, 2007). Organizations, including corporations and governments, utilize cost-benefit analyses to weigh the merits of business decisions and new policies. When conducting a cost-benefit analysis, the decision maker aims to quantify the value generated by each decision and subtract the associated costs. The option with the largest difference between benefits and costs is typically considered the best choice. The primary application of cost-benefit analyses involves informing decisions related to acceptance and facilitating evaluations. They can also be

applied to compare different options, but a separate analysis must be performed for each choice.

Risk Assessment Models

Many decisions, such as those involving monetary investments or medical procedures, are frequently approached through a formal-empirical method. In this approach, decision-makers analyze options probabilistically, aiming to minimize risk or loss, which in turn maximizes expected value. This type of decision-making often employs a gambling metaphor, where the decision-maker assesses the level of acceptable risk in relation to potential payoffs (Beach and Connelly, 2005). Contemporary decision-making theories have developed from the study of games of chance, rather than from a psychological exploration of risk or value (Tversky and Kahneman, 2000). This analysis of probabilities associated with different choices has led to research that employs a gambling metaphor, resulting in normative theories aimed at maximizing the expected outcome of any gamble. The most frequently examined methods of decision-making regard it as a process of risk assessment, where individuals consider the likelihood of mitigating risk.

Naturalistic (Descriptive) Models

Recent approaches to decision-making explore how experienced individuals make choices in complex contexts. These naturalistic methods emphasize the importance of context, identity, and unconscious emotions in the decision-making process.

Narrative-Based Decision Making

Instead of relying solely on probabilistic calculations, people often base their decisions on the explanations surrounding the available options. Decisions are rarely neutral in terms of content or domain; individuals typically have substantial knowledge about the choices before them and carry various beliefs and biases. In fact, many decisions are emotionally charged to the point where rationality plays a minimal role. This means that decisions made in different contexts are influenced by fundamental differences in how individuals subjectively evaluate each outcome, leading to varying choice strategies. These strategies may include constructing narratives, focusing on potential regret, considering moral implications, choosing favorites, avoiding the worst options, or reacting emotionally—strategies that often deviate from rationality (Jonassen, 2012).

The more personally significant a decision is, the more intricate the narrative processing involved will be. Important decisions prompt rich story construction that draws heavily on personal background knowledge (Rettinger and Hastie,

2001). Research has confirmed the vital role of story construction in the decision-making process (Pennington and Hastie, 1988).

Identity-Based Decision Making

There are many factors that influence personal and professional decisions. People often make choices based on rules associated with their identities. In other words, individuals are significantly influenced by questions related to their personal identities when making decisions. Organizations, particularly social and political ones, encourage individuals to adopt specific identities and to behave and decide in socially acceptable ways. These organizations define identities tailored to their contexts, train individuals in these identities, and socialize them to embrace them. Adherence to established identities by individuals and organizations typically results in compliance with situationally appropriate regulations. Both rules and identities serve as fundamental determinants for choices made within familial structures, economic systems, informal social gatherings, and political arenas (March, 1994). Rule-based decision-making links identities to recognized situations, making it easier for those with similar identities to make decisions. Since identities are formed through social interactions, individuals often adopt behavioral rules from families, schools, religious groups, and social or political organizations that predict their behavior and decision-making tendencies (Jonassen, 2012).

Decision-Making Process

Decision-making encompasses the process of recognizing and choosing among different options according to the values and preferences of the individual making the decision. When faced with a choice, it's crucial to evaluate several possibilities. Our objective is not just to discover as many alternatives as we can but also to select the one that most closely matches our goals, objectives, wishes, and values (Harris, 1998).

Overall, the decision-making process has the following stages (Fülop, 2005):

1-Defining the Problem

The procedure should pinpoint root causes, assumptions, limitations, interfaces, and concerns of stakeholders. The goal is to formulate a straightforward and succinct problem statement that encompasses both the current and desired conditions, although complicated issues may necessitate more elaborate statements. It is crucial for all decision-makers and stakeholders to reach a mutual understanding of this problem statement, even if achieving

consensus requires time, as this agreement is vital before proceeding (Fülop, 2005).

2-Determining Requirements

Requirements define the essential conditions that any acceptable solution to a problem must satisfy. They act as constraints for feasible solutions and should be expressed in clear, quantitative terms to ensure clarity and eliminate ambiguity. Documenting these requirements and the evaluation methods helps prevent disputes in the future.

3-Establishing Goals

Goals are objectives; they may be conflicting, but they are a concomitant of the decision-making process. Goals are comprehensive statements that reflect the wants and desires of the decision-maker.

4-Identifying Alternatives

Alternatives present diverse solutions that meet the requirements while achieving the goal.

5-Defining Criteria

Decision criteria must align with goals to effectively evaluate alternatives. Each goal should have at least one corresponding criterion, while more complex goals may require multiple criteria for accurate assessment. The criteria should be (Baker et al., 2002): Capable of distinguishing between alternatives and facilitating their performance comparison, complete with all goals, functional and significant, non-redundant, and minimal in number.

6-Selecting A Decision-Making Tool

There are numerous decision-making tools and approaches for solving different types of decision-making problems. Some of them will be mentioned in this book chapter. The mentioned ones were chosen based on their convenience with the forestry-related situations. Selecting the right tool/approach is not an easy task; it depends on the problem's structure and the decision maker's needs. We do have some simpler and more complex methods. They are not inferior or superior to each other. However, complex problems need complex methods. Multi-Criteria Decision-Making (MCDM) tools are primarily used for complex decision-making.

7-Evaluating the Alternatives

Every effective decision-making method relies on input data to assess alternatives against defined criteria. This evaluation can be either objective, based on universally understood scales of measurement (such as monetary value), or subjective, reflecting the evaluator's personal judgment and interpretation (Fülop, 2005). Once these assessments are completed, an appropriate decision-making tool/approach can be employed for ranking alternatives.

8- Validating the Solutions

The selected alternatives and the overall solution should always be validated. There is always a possibility of misapplication of the decision-making tool. In complex models, these validation processes are crucial. In general, for validation, different decision-making tools are used separately for the same decision problem, and the results are compared.

Multi-Criteria Decision-Making Methods

Multi-Criteria Decision Making (MCDM) techniques provide a versatile tool for managing various variables evaluated through diverse means, aiding decision-makers in clarifying complex issues (Dhanalakshmi et al., 2022). When there are multiple criteria and alternatives, the MCDM methods are employed to decide the most appropriate alternative (Ersoy, 2021). In the real world, alternatives for solutions are mostly evaluated against multiple conflicting criteria, which is a common situation that necessitates the use of MCDM (Ghorabaee et al., 2016). MCDM methods provide an effective and organized approach to addressing decision-making challenges. By employing these techniques, you can confidently and clearly manage complex choices (Sahoo and Goswami, 2023).

In recent years, significant advancements have been made in MCDM approaches. Traditional methods, such as the Analytic Hierarchy Process, TOPSIS, and ELECTRE, have established a solid foundation for effective decision-making (Alvarez et al., 2021). However, new methodologies have emerged that enhance these traditional approaches. These include multi-objective MCDM methods that tackle conflicting objectives, fuzzy-based approaches that handle uncertainty and imprecision, data-driven models that leverage machine learning and big data analytics, and hybrid methodologies that combine multiple methods (Sahoo and Goswami, 2023). MCDM methods have been widely applied across multiple domains. From environmental sciences to business management, healthcare to public policy (Khan et al., 2021).

Multi-Criteria Decision-Making (MCDM) methods are vital for addressing the challenges associated with decisions that encompass numerous goals,

evaluation criteria, and stakeholder interests (Hasan et al., 2022). These methods establish a systematic and organized decision framework, allowing decision-makers to structure complex issues into well-defined criteria, assess various alternatives, and make rational choices guided by transparent and consistent principles (Li et al., 2020).

A major advantage of MCDM is its capacity to consider multiple, and often competing, objectives simultaneously. By explicitly modeling trade-offs and preferences among these objectives, MCDM enables the identification of solutions that align with the decision-makers' priorities and strategic goals (Sahoo and Goswami, 2024; Gebre et al., 2021). Moreover, these approaches are well-equipped to manage uncertainty and subjective judgment, employing analytical tools such as fuzzy logic and probabilistic modeling to support sound decisions even when information is incomplete or imprecise (Pelissari et al., 2022).

Another key contribution of MCDM is its ability to incorporate diverse stakeholder perspectives by systematically integrating their criteria and preferences into the evaluation process (Yenugula et al., 2024). This participatory dimension fosters inclusiveness, equity, and stakeholder acceptance, thereby enhancing the legitimacy and credibility of the final decisions.

Due to their flexibility and broad applicability, MCDM methods are employed in a wide range of disciplines, including business, engineering, healthcare, environmental management, and public policy (Supriya and Gadekallu, 2023). They support decision-making in various contexts such as project prioritization, strategic planning, risk assessment, and resource allocation.

Overall, MCDM techniques significantly enhance decision quality by providing a comprehensive, transparent, and participatory framework for decision-making. Through the integration of multiple objectives, criteria, and stakeholder viewpoints, they facilitate more informed, balanced, and equitable outcomes in complex and uncertain environments (Vatankhah et al., 2023).

There are MCDM methods that have been widely applied in various fields. For example, the Analytic Hierarchy Process (AHP) is commonly used in business and management decisions for prioritizing objectives and evaluating alternative strategies. The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) has been utilized to rank the suppliers. ELECTRE: ÉLimination Et Choix Traduisant la REalité ("Elimination and Choice Translating Reality") has been employed to evaluate investment alternatives. Fuzzy Logic-based MCDM methods, such as Fuzzy AHP and Fuzzy TOPSIS, have been utilized in engineering and technology applications for product design, as well as alternative evaluation regarding cost, reliability, and performance. The Preference Ranking Organization METHod for Enrichment of Evaluations

(PROMETHEE), TOPSIS, and ELECTRE have been employed in project management to rank project activities and distribute resources. In environmental decision-making, PROMETHEE, AHP, and TOPSIS have been utilized (Sahoo and Goswami, 2023).

Decision Support Systems

The idea of computer-aided decisions dates back to 1963 (Bonini, 1963). The term "decision support systems" (DSS) was first used in the early 1970s (Eom and Kim, 2006). Since then, researchers have defined DSS in various ways. Jones (1980) characterized a DSS as a computer-based tool designed to assist decision-makers in addressing semi-structured problems and enhancing decision quality. According to Sheng and Zhang (2009), a DSS is a human-computer interactive system capable of gathering, processing, and delivering information through computational methods. Yazdani et al. (2017) described it as a distinct category of computerized information systems that facilitate and manage decision-making processes. Similarly, Terribile et al. (2015) defined it as an intelligent system that offers operational solutions and supports decision-making by utilizing collected data to address specific needs and challenges.

Since the 1980s, the scope of DSS applications has been expanded due to the contributions of universities and various organizations (Power, 2008). These contributions paved the way for DSS to extend its use beyond regular business and management applications. Today, DSS is a fruitful area for both research and development (Liu et al., 2010). Numerous decision support systems (DSS) were created to assist decision-makers at every level within an organization, encompassing systems that facilitate problem structuring, operations, financial management, and strategic decision-making, as well as extending to support optimization and simulation (Liu et al., 2010). Although more powerful complementary support tools than traditional DSS have been developed over time, such as Enterprise Information Systems (EIS) and Group Decision Support Systems (GDSS), they remain less prevalent. GDSS is utilized to support team problem-solving and group decision-making by employing brainstorming, idea evaluation, and communication resources. EIS is suitable for personal or small group use at the corporate level.EIS can provide numerous pieces of information, such as success metrics, key indicators, reports, and budget details (Anson et al., 1995; Costa et al., 2003; DeSanctis and Gallupe, 1987; Limayem et al., 2006; Elam and Leidner, 1995; Kosata, 1994; Power, 2003).

Due to the new challenges and complex decisions that need to be made, the traditional DSS must be improved. Today, both researchers and practitioners

consistently work on making better decisions, and new methods or approaches have been developed.

The integration of Multi-Criteria Decision-Making (MCDM) methods with Decision Support Systems (DSS) is mutually advantageous. MCDM techniques assist DSS by identifying, analyzing, and comparing competing alternatives, whereas DSS platforms provide an effective means to implement MCDM methodologies and to store, manage, and retrieve MCDM models. Over time, MCDM has played a significant role in advancing the development of various DSS subfields and specialized applications (Kou et al., 2011). DSS were conceived to augment MCDM methodologies, consequently streamlining the process of decision making. These systems utilize sophisticated technologies and applications to empower decision-makers in the efficient exploitation of data through diverse modeling strategies, thereby enabling the resolution of partially structured and ill-defined issues (Qian et al., 2004).

Various DSS Applications in Forestry

In this subtitle of the chapter, examples from the current literature will be provided. The main intention of this part is to prove to the readers how you can implement and benefit from various MCDM methods in forestry. Hopefully, readers can embrace the knowledge and conduct their own original research.

In 1981, Hall described an approach to forest management decision-making (Hall, 1981). He acknowledged subjective and objective elements to offer an encouraging methodology for more creative design in forest planning. Varma and friends tried to create a DSS for sustainable forest management (Varma et al., 2000). In 2006, Blennow and Salinas (Blennow and Salinas, 2006) utilized DSS for active risk management in sustainable forestry. Ananda and Herath published a comprehensive review of MCDM methods and their application in forest management planning in 2009 (Ananda and Herath, 2009). Lawrence and Stewart discussed the success of the participatory forest decision-making process, aided by computer-based tools (Lawrence and Stewart, 2011).

In 2012, Marcot and friends published a paper on recent advances in applying decision science to manage national forests in the USA (Marcot et al., 2012). Another comprehensive review of decision-making methods to handle uncertainty and risk in forest management under climate change was conducted by Yousefpour and colleagues in 2012 (Yousefpour et al., 2012). Hemming and colleagues tried to conceptualize biodiversity conservation and decision science in 2022 (Hemming et al., 2022). Güngör and Sen employed the AHP method for selecting suitable forest areas for honey production (Güngör and Sen, 2018). Lakicevic and friends utilized AHP for smarter forestry decision making

(Lakicevic et al., 2018). Ortiz-Urbina and colleagues conducted a review of Multi-Criteria Group Decision Making Methods in forestry (Ortiz-Urbina et al., 2019). In 2020, Bayram adopted a hybrid approach of the Entropy and TOPSIS methods to evaluate the economic contribution of forest products in Turkiye (Bayram, 2020). In 2021, Bayram employed a Fuzzy AHP approach to evaluate sustainable forest management criteria and indicators in Turkiye (Bayram, 2021). In 2023, Matilainen and colleagues conducted a comprehensive review of the decision-making process of private forest owners and forest-based services (Matilainen et al., 2023). Gungor and Sen employed a hybrid MCDM model for Sustainable Afforestation Strategies in post-mining rehabilitation (Güngör and Şen, 2024).

It can be said that, based on the given examples, MCDM models and DSS are primarily employed in forest management-related issues due to their inherent nature and complexity. However, the forest products industry and product-based problems can also be a fruitful topic for adopting decision science and employing MCDM methods.

References

Alvarez, P. A., Ishizaka, A., & Martinez, L. (2021). Multiple-criteria decision-making sorting methods: A survey. Expert Systems with Applications, 183, 115368. https://doi.org/10.1016/j.eswa.2021.115368

Ananda, J., & Herath, G. (2009). A critical review of multi-criteria decision making methods with special reference to forest management and planning. Ecological Economics, 68(10), 2535–2548. https://doi.org/10.1016/j.ecolecon.2009.05.010

Anson R, Bostrom RP, Wynne BE (1995) An experiment assessing group system and facilitator effects on meeting outcomes. Manag Sci 41(2):189–208

Armstrong, J. S. (1982). The value of formal planning for strategic decisions. Strategic Management Journal, 3(3), 197–211.

Baker, D., Bridges, D., Hunter, R., Johnson, G., Krupa, J., Murphy, J. and Sorenson, K. (2002) Guidebook to DecisionMaking Methods, WSRC-IM-2002-00002, Department of Energy, USA.

Bayram, B. Ç. (2020). Evaluation of forest products trade economic contribution by entropy-TOPSIS: Case study of Turkey. BioResources, 15(1), 1419–1429. https://doi.org/10.15376/biores.15.1.1419-1429

Bayram, B. Ç. (2021). A sustainable forest management criteria and indicators assessment using fuzzy analytic hierarchy process. Environmental Monitoring and Assessment, 193(7), 425. https://doi.org/10.1007/s10661-021-09176-x

Beach, L. R., & Connelly, T. (2005). The psychology of decision making: People in organizations (2nd ed.). Thousand Oaks, CA: Sage.

Blennow, K., & Salinas, O. (2006). Decision Support for Active Risk Management in Sustainable Forestry. Journal of Sustainable Forestry, 21(2–3), 201–212. https://doi.org/10.1300/J091v21n02_12

Bonini CP (1963). Simulation of Information and Decision Systems in the Firm. Prentice Hall: Englewood Cliffs." (Eom and Kim, 2006, p. 1275)

Costa JP, Melo P, Godinho P, Dias LC (2003) The AGAP system: a GDSS for project analysis and evaluation. Eur J Oper Res 145(2):356–367

DeSanctis G, Gallupe RB (1987) A foundation for the study of group decision support systems. Manag Sci 33(5):589–609

Dhanalakshmi, C. S., Madhu, P., Karthick, A., Mathew, M., & Vignesh Kumar, R. (2022). A comprehensive MCDM-based approach using TOPSIS and EDAS as an auxiliary tool for pyrolysis material selection and its application. Biomass Conversion and Biorefinery, 12(12), 5845–5860. https://doi.org/10.1007/s13399-020-01009-0

Elam JJ, Leidner DG (1995) EIS adoption, use and impact: the executive perspective. Decis Support Syst 14(2):89–103

Eom, S., & Kim, E. (2006). A survey of decision support system applications (1995–2001). Journal of the Operational Research Society, 57(11), 1264–1278. https://doi.org/10.1057/palgrave.jors.2602140

Ersoy, Y. (2021). Equipment selection for an e-commerce company using Entropy-based TOPSIS, EDAS and CODAS methods during the COVID-19. Logforum, 17(3), 341–358. https://doi.org/10.17270/J.LOG.2021.603

Fülop, J. (2005). Introduction to decision making methods. BDEI-3 Workshop, Washington.

Gebre, S. L., Cattrysse, D., Alemayehu, E., & Van Orshoven, J. (2021). Multi-criteria decision making methods to address rural land allocation problems: A systematic review. International Soil and Water Conservation Research, 9(4), 490-501. https://doi.org/10.1016/j.iswcr.2021.04.005

Ghorabaee, M. K., Zavadskas, E. K., Amiri, M., & Turskis, Z. (2016). Extended EDAS Method for Fuzzy Multi-criteria Decision-making: An Application to Supplier Selection. International Journal Of Computers Communications & Control, 11(3), Article 3. https://doi.org/10.15837/ijccc.2016.3.2557

Güngör, E., & Sen, G. (2018). Selecting Suitable Forest Areas For Honey Production Using The Ahp: A Case Study In Turkey. Cerne, 24, 67–79. https://doi.org/10.1590/01047760201824012511

Güngör, E., & Şen, G. (2024). Sustainable Afforestation Strategies: Hybrid Multi-Criteria Decision-Making Model in Post-Mining Rehabilitation. Forests, 15(5), 783. https://doi.org/10.3390/f15050783

Hall, T. H. (1981). Forest Management Decision-Making: Art or Science? The Forestry Chronicle, 57(5), 233–238. https://doi.org/10.5558/tfc57233-5

Harris, R. (1998) Introduction to Decision Making, VirtualSalt. http://www.virtualsalt.com/crebook5.htm

Hasan, M. G., Ashraf, Z., & Khan, M. F. (2022). Multi-choice best-worst multi-criteria decision-making method and its applications. International Journal of Intelligent Systems, 37(2), 1129-1156. https://doi.org/10.1002/int.22663

Hemming, V., Camaclang, A. E., Adams, M. S., Burgman, M., Carbeck, K., Carwardine, J., Chadès, I., Chalifour, L., Converse, S. J., Davidson, L. N. K., Garrard, G. E., Finn, R., Fleri, J. R., Huard, J., Mayfield, H. J., Madden, E. M., Naujokaitis-Lewis, I., Possingham, H. P., Rumpff, L., … Martin, T. G. (2022). An introduction to decision science for conservation. Conservation Biology, 36(1), e13868. https://doi.org/10.1111/cobi.13868

Jonassen, D. H. (2000). Toward a design theory of problem solving. Educational Technology Research and Development, 48(4), 63–85.

Jonassen, D. H. (2012). Designing for decision making. Educational Technology Research and Development, 60(2), 341–359. https://doi.org/10.1007/s11423-011-9230-5

Jones, J.W., 1980. Decision support systems – An organizational perspective – KEEN, PG, MORTON. MSS. Adm. Sci. Q. 25 (2), 376–382. https://doi.org/10.2307/2392463.

Khan, S. A. R., Yu, Z., Golpira, H., Sharif, A., & Mardani, A. (2021). A state-of-the-art review and meta-analysis on sustainable supply chain management: Future research directions. Journal of Cleaner Production, 278, 123357. https://doi.org/10.1016/j.jclepro.2020.123357

Klein, G. (1998). Sources of power: How people make decisions. Cambridge, MA: MIT Press.

Kosaka T (1994) The first investigation of executive information systems practices in Japanese firms. In: Proceedings of the TIMS/ORMA conference, Boston

Kou, G., Shi, Y., & Wang, S. (2011). Multiple criteria decision making and decision support systems—Guest editor's introduction. Decision Support Systems, 51(2), 247–249. https://doi.org/10.1016/j.dss.2010.11.027

Lakicevic, M., Srdjevic, B., & Velichkov, I. (2018). Combining AHP and Smarter in Forestry Deci- sion Making. BALTIC FORESTRY, 24(1).

Lawrence, A., & Stewart, A. (2011). Sustainable forestry decisions: On the interface between technology and participation. Mathematical and Computational Forestry & Natural-Resource Sciences (MCFNS), 3(1), 42-52 (11).

Lewin, K. (1943). Defining the ''Field at a given time''. Psychological Review, 50, 292–310.

Li, T., Li, A., & Guo, X. (2020). The sustainable development-oriented development and utilization of renewable energy industry-A comprehensive analysis of MCDM methods. Energy, 212, 118694. https://doi.org/10.1016/j.energy.2020.118694

Limayem M, Banerjee P, Ma L (2006) Impact of GDSS: opening the black box. Decis Support Syst 42:945–957

Liu, S., Duffy, A. H. B., Whitfield, R. I., & Boyle, I. M. (2010). Integration of decision support systems to improve decision support performance. Knowledge and Information Systems, 22(3), 261–286. https://doi.org/10.1007/s10115-009-0192-4

Marcot, B. G., Thompson, M. P., Runge, M. C., Thompson, F. R., McNulty, S., Cleaves, D., Tomosy, M., Fisher, L. A., & Bliss, A. (2012). Recent advances in applying decision science to managing national forests. Forest Ecology and Management, 285, 123–132. https://doi.org/10.1016/j.foreco.2012.08.024

Matilainen, A., Andersson, E., Lähdesmäki, M., Lidestav, G., & Kurki, S. (2023). Services for What and for Whom? A Literature Review of Private Forest Owners' Decision-Making in Relation to Forest-Based Services. Small-Scale Forestry, 22(3), 511–535. https://doi.org/10.1007/s11842-023-09541-3

Mishan, E. J., & Quah, E. (2007). Cost benefit analysis (5th ed.). New York: Routledge.

Ortiz-Urbina, E., González-Pachón, J., & Diaz-Balteiro, L. (2019). Decision-Making in Forestry: A Review of the Hybridisation of Multiple Criteria and Group Decision-Making Methods. Forests, 10(5), 375. https://doi.org/10.3390/f10050375

Pelissari, R., Khan, S. A., & Ben-Amor, S. (2022). Application of Multi-Criteria Decision-Making Methods in Sustainable Manufacturing Management: A Systematic Literature Review and Analysis of the Prospects. International Journal of Information Technology & Decision Making, 21(02), 493-515. https://doi.org/10.1142/S0219622021300020

Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(3), 521–533.

Power D.J. (2003) A brief history of decision support systems. DSS Resource.COM, available at http:// DSSResources.COM/history/dsshistory.html

Power, D. J. (2008). Decision Support Systems: A Historical Overview. In F. Burstein & C. W. Holsapple, Handbook on Decision Support Systems 1 (pp. 121–140). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-48713-5_7

Qian Z, Huang GH, Chan CW. 2004. Development of an intelligent decision support system for air pollution control at coal-fired power plants. Expert System with Applications 26(3): 335–356.

Rettinger, D. A., & Hastie, R. (2001). Content effects on decision making. Organizational Behavior and Human Decision Process, 85(2), 336–359.

Sahoo, S. K., & Goswami, S. S. (2023). A Comprehensive Review of Multiple Criteria Decision-Making (MCDM) Methods: Advancements,

Applications, and Future Directions. Decision Making Advances, 1(1), 25–48. https://doi.org/10.31181/dma1120237

Sahoo, S., & Goswami, S. (2024). Theoretical framework for assessing the economic and environmental impact of water pollution: A detailed study on sustainable development of India. Journal of Future Sustainability, 4(1), 2334. http://dx.doi.org/10.5267/j.jfs.2024.1.003

Sheng, Y.K., Zhang, S., 2009. Analysis of problems and trends of decision support systems development. In: In: 2009 International Conference on E-Business and Information System Security, pp. 1216–1218.

Supriya, Y., & Gadekallu, T. R. (2023). A Survey on Soft Computing Techniques for Federated Learning-Applications, Challenges and Future Directions. ACM Journal of Data and Information Quality. https://doi.org/10.1145/3575810

Terribile, F., Agrillo, A., Bonfante, A., Buscemi, G., Colandrea, M., D'Antonio, A., De Mascellis, R., De Michele, C., Langella, G., Manna, P., Marotta, L., Mileti, F.A., Minieri, L., Orefice, N., Valentini, S., Vingiani, S., Basile, A., 2015. A web-based spatial decision supporting system for land management and soil conservation. Solid Earth. 6, 903–928. https://doi.org/10.5194/se-6-903-2015.

Tversky, A., & Kahnemann, D. (2000). Rational choice and the framing of decisions. In D. Kahnemann & A. Tversky (Eds.), Choices, values, and frames (pp. 209–223). Cambridge: Cambridge University Press.

Varma, V. K., Ferguson, I., & Wild, I. (2000). Decision support system for the sustainable forest management. Forest Ecology and Management, 128(1–2), 49–55. https://doi.org/10.1016/S0378-1127(99)00271-6

Vatankhah, S., Darvishmotevali, M., Rahimi, R., Jamali, S. M., & Ale Ebrahim, N. (2023). Assessing the application of multi-criteria decision making techniques in hospitality and tourism research: a bibliometric study. International Journal of Contemporary Hospitality Management. https://doi.org/10.1108/IJCHM-05-2022-0643

Yates, J. F. (2003). Decision management: How to assure better decisions in your company. San Francisco, CA: Joissey-Bass.

Yates, J. F., & Tschirhart, M. D. (2006). Decision-making expertise. In K. A. Ericsson, N. Charness, P. F. Feltovich, & R. R. Hoffman (Eds.), The Cambridge handbook of expertise and expert performance (pp. 421–438). Cambridge: Cambridge University Press.

Yazdani, M., Zarate, P., Coulibaly, A., Zavadskas, E.K., 2017. A group decision making support system in logistics and supply chain management. Expert Syst. Appl. 88, 376–392. https://doi.org/10.1016/j.eswa.2017.07.014.

Yenugula, M., Sahoo, S., & Goswami, S. (2024). Cloud computing for sustainable development: An analysis of environmental, economic and social benefits. Journal of Future Sustainability, 4(1), 59-66. http://dx.doi.org/10.5267/j.jfs.2024.1.005

Yousefpour, R., Jacobsen, J. B., Thorsen, B. J., Meilby, H., Hanewinkel, M., & Oehler, K. (2012). A review of decision-making approaches to handle uncertainty and risk in adaptive forest management under climate change. Annals of Forest Science, 69(1), 1–15. https://doi.org/10.1007/s13595-011-0153-4

Chapter 10

Application of the AHP Approach in Forest Road Type Selection

Sedef Seda GÜLCÜ[1], Korhan ENEZ[2*]

Introduction

Forest roads are the essential infrastructure required for the performance of technical and economic analyses in forestry. The building and maintenance of forest roads significantly alter the ecological structure of nature, making the decisions taken during the design phase and their subsequent implications very important. The development of forest roads is an irreversible technical endeavor (Erdaş, 1997; Bayoğlu, 1997; Erdaş et al., 2014). Moreover, forestry activities involve the administration of human endeavors to satisfy the permanent demands of forests with limited resources (Erdaş et al., 2014). The management of forests, which support 80% of terrestrial biodiversity, safeguard resources like air, water, and soil, and contain many plants, animals, and organisms, can only be regarded as successful if their sustainability is guaranteed (FAO, 2023). Sustainable forest management is the utilization and regulation of various forest types at local, national, and international levels, aimed at preserving their ecological diversity, productivity, and self-renewal capacity, both biologically and economically, without compromising other ecological systems (GDF, 2019).

In the context of sustainable management, all regulations and measures designed to access a particular forest area, transport requisite production equipment to that location, and extract forest products while minimizing potential harm to the forest soil, stands, intended products, and natural landscape for the benefit of society are referred to as "opening for harvesting" (Erdaş et al., 2014). This indicates that forest roads are the most essential and critical infrastructure for facilitating forest operations. Consequently, it is essential to prioritize this significance from the outset of the decision-making process. The significance of identifying the type of forest road is once again highlighted at this point in time.

[1] Forestry Engineering, 37150, Kastamonu, Türkiye, sedefsedag@gmail.com
[2] Kastamonu University, Faculty of Forestry, Department of Forestry Engineering, Kastamonu, Türkiye, Orcid: 0000-0001-7526-0032, korhanenez@kastamonu.edu.tr

The total requirement for forest roads in the country is 210,000 kilometers. Out of this total, 140,000 km have been completed (Acar, 2005). Data from the General Directorate of Forestry (GDF) indicates that 24,330 km of new roads were constructed from 2005 to 2021 (GDF, 2023), with 45,670 km of road still to be developed. Moreover, an analysis of data from the General Directorate of Forestry indicates that new road development averaged almost 1,000 kilometers per year from 1998 to 2021. Consequently, we might infer that additional roads will be constructed over the following 45 years. Despite a projected road density of 20 m/ha upon road completion, and the modified road density indicator from the 2008 circular, which is deemed to constitute no more than 1% of the forested area, the average road density in European nations stands at 30 m/ha. Acknowledging possible differences in calculations, it is evident that decisions pertaining to road development will significantly influence our future. We expect that the study's conclusions will have long-lasting implications, allowing us to make informed decisions for the future. The study's methodological findings are expected to provide a distinctive contribution to forestry methodology.

Material and Methods

The study was carried out in the Kastamonu Regional Directorate of Forestry. The study involved a review of the literature. Expert interviews were performed, and expert opinions were considered, in conjunction with a prepared survey. The Analytical Hierarchy Process (AHP) was initially introduced by Myers and Alpet in 1968 and further refined by Saaty in 1977 as a technique for addressing decision-making challenges (Saaty, 1980). The Analytic Hierarchy Process (AHP) is a multi-criteria decision-making methodology that assesses both quantitative and qualitative criteria, incorporates the preferences, experiences, intuition, knowledge, judgments, and insights of individuals or groups, and facilitates the resolution of complex problems through a hierarchical framework. The decision-maker can integrate both objective and subjective considerations into the decision-making process. Consequently, this affords the decision-maker the ability to comprehend their own decision-making processes.

AHP is a hierarchical structure that includes a minimum of three layers. The highest level of the hierarchy is the objective. The subsequent level comprises the primary criteria and, if relevant, sub-criteria. The lowest level comprises decision alternatives. For pairwise comparisons to maintain consistency, the criteria must be precisely quantified, and each criterion must be well delineated. Criteria ought to be categorized according to their shared attributes. The Analytic Hierarchy Process (AHP) is applicable to numerous criteria and serves as an effective tool for collective decision-making. Sensitivity analysis facilitates the examination of

the result's flexibility. The establishment of the hierarchy and pairwise comparison matrices is inherently subjective, necessitating the involvement of experienced and expert personnel.

The algorithmic steps of the AHP method are as follows.

Step 1: The issue is delineated. The necessary criteria for the decision are established, and the priorities of those criteria are defined.

Step 2: A hierarchical framework is established. The primary objective is positioned at the highest point. The primary criteria and their corresponding sub-criteria are listed below. The alternatives occupy the lowest level of the hierarchy. The quantity of tiers in the hierarchy is contingent upon the complexity and specificity of the issue. In establishing the hierarchy, it is presumed that alternatives at the same level are entirely independent of one another.

Step 3: A pairwise comparison matrix is constructed. Employing a significance scale from 1 to 9, matrices are constructed to evaluate decision alternatives based on the primary criterion, subsequently addressing any sub-criteria, and ultimately encompassing all criteria. Comparison matrices are square matrices characterized by diagonal elements equal to 1 (Mu and Pereyra-Rojas, 2017).

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \vdots & \vdots & \dots & \dots & \vdots \\ \vdots & \vdots & \dots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{bmatrix}$$

Decision alternatives are evaluated independently against each criterion. Decision matrices are built with the 1-9 comparison scale proposed by Saaty (1980) as illustrated in Table 1.

Table 1. Comparison scale

Importance Scale	Definitions	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Moderate importance	One activity is slightly favored over another
5	Strong importance	One activity is strongly favored over another
7	Very Strong importance	One activity is very strongly favored over another
9	Extreme importance	One activity is of the highest possible order of affirmation
2, 4, 6, 8	Intermediate values	When compromise is needed

Step 4: The pairwise comparison matrices perform normalization. Each element in the matrix is normalized by dividing it by the sum of its column. The total of each column in the normalized matrix equals 1.

$$a_{ij}' = \frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}} \quad , \quad i,j = 1,2,\ldots,n$$

Step 5: The priority vector is computed. The sum of each row of the normalized matrix is divided by the matrix size to compute the average. These values represent the significance weights computed for each criterion. These weights constitute the priority vector..

$$w_i = \left(\frac{1}{n}\right) \sum_{j=1}^{n} a_{ij}' \quad , \quad i,j = 1,2,\ldots,n$$

Consequently, percentage importance distributions are derived, illustrating the relative significance of the criteria.

Four distinct methodologies may be employed in the formulation of the priority vector. These can be categorized as the Most General Method, an Improved Method, a Satisfactory Method, and the Optimal Method. When the comparison matrices exhibit consistency, all four approaches will produce identical outcomes.

Step 6: The consistency ratio is computed. Following the execution of pairwise comparisons and the establishment of their priorities, the consistency of the comparison matrices is assessed. To assess the consistency of a matrix A derived from pairwise comparisons, it is necessary to compute the "Consistency Index (CI)," which is one of several methodologies available. The CI coefficient is determined using the formula (Chen et al., 2010; Saaty, 2001). To evaluate consistency, the value of the "Random Index (RI)" must be established. The RI values established for n-dimensional comparison matrices are presented in Table 2 (Saaty, 1980).

Table 2. RI values according to the dimensions of the comparison matrices

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.53	1.56	1.57	1.59

If the CR is less than 0.10, it is decided that the comparison matrix is consistent (Saaty, 2012).

$$CI = \lambda_{max} - n/n - 1$$
$$CR = CI/RI$$

Step 7: A pairwise comparison matrix is constructed for the criteria, and the priority vector for the choice alternatives is computed. The priority vector can also be characterized as the weight vector for the criteria.

Step 8: The alternatives are prioritized. The priority vectors derived from the criteria are amalgamated to produce the comprehensive priority matrix. The resultant vector is derived by multiplying the priority matrix by the priority vectors of the decision alternatives and summing them. The decision option with the greatest weight in this vector is identified as the preferred choice for resolving the issue.

Eleven distinct characteristics were employed in the selection of the forest road type. The characteristics include vehicle type (VT), road width (RW), amount of timber (AT), vehicle loading capacity (VL), vehicle speed (VS), slope (SL), slope direction (SD), number of vehicles (NV), soil type (ST), forest function (FF), and geological structure (GS).

Results

The pairwise comparison resulting from the examination is shown in tabular format below (Table 3). This 11 x 11 matrix study utilized a constant value of 1.51 for the computation of the CR value.

Table 3. Comparison matrix of the factors used in the study

Item Description	VT	RW	AT	VL	VS	SL	SD	NV	DT	FF	GS
VT	1.00	1.00	2.00	2.00	2.00	3.00	3.00	3.00	4.00	4.00	5.00
RW	1.00	1.00	1.00	1.00	2.00	3.00	3.00	3.00	4.00	5.00	6.00
AT	0.50	1.00	1.00	2.00	2.00	3.00	3.00	3.00	4.00	5.00	7.00
VL	0.50	1.00	0.50	1.00	1.00	3.00	3.00	5.00	3.00	4.00	9.00
VS	0.50	0.50	0.50	1.00	1.00	1.00	3.00	4.00	5.00	5.00	8.00
SL	0.33	0.33	0.33	0.33	1.00	1.00	2.00	2.00	2.00	4.00	6.00
SD	0.33	0.33	0.33	0.33	0.33	0.50	1.00	2.00	1.00	4.00	6.00
NV	0.33	0.33	0.33	0.20	0.25	0.50	0.50	1.00	1.00	2.00	3.00
ST	0.25	0.25	0.25	0.33	0.20	0.50	1.00	1.00	1.00	3.00	5.00
FF	0.25	0.20	0.20	0.25	0.20	0.25	0.25	0.50	0.33	1.00	5.00
GS	0.20	0.17	0.14	0.11	0.13	0.17	0.17	0.33	0.20	0.20	1.00

The findings indicate that vehicle type is the predominant influence in road type choices (Table 4). The subsequent factors were crop yield and road width, but the least relevant factor was geological structure. Alongside AHP, a multi-criteria decision support system, alternative decision-making methodologies should also be employed in the selection of forest road types.

Table 4. Ranking of parameters according to coefficient values and consistency rates

Parameters	Coefficient	CR
VT	0.1724	
RW	0.1520	
AT	0.1553	
VL	0.1331	
VS	0.1187	
SL	0.0759	0,017
SD	0.0589	
NV	0.0415	
ST	0.0463	
FF	0.0300	
GS	0.0159	

Discussion

The Analytic Hierarchy Process (AHP) is a method that offers a systematic approach to multi-criteria decision-making (di Semenanjung Malaysia, 2012), effectively addresses intricate decision problems that involve qualitative assessments (Saaty 1980), and is widely utilized in various decision-making applications (Vaidya and Kumar, 2006) due to its capacity to handle a range of multi-dimensional, immeasurable, and incomparable objectives (Ananda and Herath, 2003). In forestry, it is utilized for forest resource planning (Kangas, 1992), security risk assessments in wood raw material production (Ünver and Ergenç, 2021), implementation of forest certification (Alves et al., 2019), identification of stands vulnerable to forest pests (Sivrikaya et al., 2022; Sivrikaya et al., 2023) and landslide-prone areas (Aksoy,

2023; Thammaboribal et al., 2025), planning of forest roads (Mohammadi Samani et al., 2010; Pellegrini and Grigolato, 2013), and in various interdisciplinary studies.

Forest road planning considers physical, economic, and environmental criteria (Naghdi, Mohammadi Limaei, 2009). Due to the high costs associated with forest road building (Mohammadi Saman, 2010), the economics of transporting wood raw materials are crucial for planning (di Semenanjung Malaysia, 2012). The objective is to evaluate the capacity of forest roads based on their many roles in accordance with forestry goals and to strategize their future usage conditions (Gümüş, 2009). The analysis revealed that vehicle type is the most significant determinant in road type choosing. The subsequent factors were crop yield and road width, whilst geological structure was the least relevant factor. Transportation methods in forestry may differ based on topographical conditions, especially ground gradient. Nonetheless, inappropriate machinery selection and unanticipated operations may result in adverse ecological consequences, including soil degradation (Akay and Yılmaz, 2017).

In their study, di Semenanjung Malaysia (2012) found slope as the principal factor influencing suitable forest road allocation through paired comparison matrices. Akay et al. (2018) examined the risks associated with forest road design and construction with the Fuzzy Analytical Hierarchy Process (FAHP) methodology. The research identified the primary sub-risk factors as deficient road planning, insufficient on-site safety protocols, capital shortages, legal complications pertaining to road planning, landslide hazards during construction, and illicit deforestation.

This study employs the AHP technique, a multi-criteria decision support system, for selecting forest road types, thereby considerably enhancing scientifically informed decision-making in forest road planning procedures. Such methodologies provide a comprehensive assessment of forest roads from economic, environmental, technical, and social viewpoints. Future research should incorporate additional multi-criteria decision-making methodologies alongside the AHP method to enhance the reliability and validity of the findings. Moreover, the amalgamation of these methodologies would provide a comparative assessment of the environmental repercussions, economic viability, and operational efficacy of forest road development across various scenarios. This will facilitate decisions in forest road planning that are both technically viable and aligned with sustainable resource management, ecosystem preservation, and long-term forestry goals.

Acknowledge

This research was funded by the TÜBİTAK 2209-A University Students Research Projects Support Program.

References

Acar, H. H. (2005). Orman Yolları, Ders Notları. Karadeniz Teknik Üniversitesi, Orman Fakültesi, Ders Teksirleri Serisi: 82, 183s., Trabzon

Akay, A. E., Yılmaz, B. (2017). Using GIS and AHP for planning the primary transportation of forest products. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *4*, 19-24.

Akay, A. O., Demir, M., Akgul, M. (2018). Assessment of risk factors in forest road design and construction activities with fuzzy analytic hierarchy process approach in Turkey. *Environmental monitoring and assessment*, *190*(9), 561.

Aksoy, H. (2023). Determination of landslide susceptibility with Analytic Hierarchy Process (AHP) and the role of forest ecosystem services on landslide susceptibility. *Environmental Monitoring and Assessment*, *195*(12), 1525.

Alves, R. R., Fraj-Andrés, E., Rojo-Alboreca, A., Gracioli, C. R. (2019). Implementation of forest certification in Brazil, Spain and Portugal: an Analytic Hierarchy Process (AHP) application. *International Forestry Review*, *21*(1), 11-22.

Ananda, J., Herath, G. (2003). The use of Analytic Hierarchy Process to incorporate stakeholder preferences into regional forest planning. *Forest policy and economics*, *5*(1), 13-26.

Chen, Y., Yu, J., Khan, S., (2010). Spatial sensitivity analysis of multicriteria weights in GIS-based land suitability evaluation. Environ. Model. Softw. 25 (12), 1582–1591.

di Semenanjung Malaysia, P.H.A.A. (2012). Developing Priorities and Ranking for Suitable Forest Road Allocation using Analytic Hierarchy Process (AHP) in Peninsular Malaysia.*Sains Malaysiana*, *41*(10), 1177-1185.

Erdaş , O., (1997). Orman Yolları Cilt-I, KTÜ Orman Fakültesi Yayınları No: 187/25, Trabzon.

Erdaş, O., Acar, H.H., Eker, M., (2014). Orman Ürünleri Transport Teknikleri, Karadeniz Teknik Üniversitesi, Orman Fakültesi, Genel Yayın No: 233, Fakülte Yayın No: 39, 504s., Trabzon

GDF, Orman Genel Müdürlüğü, (2008). Orman yolları planlaması, yapımı ve bakımı, Orman Genel Müdürlüğü, Tebliğ No 292, Ankara.

Gümüş, S. (2009). Constitution of the forest road evaluation form for Turkish forestry. African Journal of Biotechnology. 8, 20.

Kangas, J. (1992). Multiple-use planning of forest resources by using the analytic hierarchy process. *Scandinavian Journal of Forest Research*, *7*(1-4), 259-268.

Mohammadi Samani, K., Hosseini, S. A., Lotfian, M., Najafi, A. (2010). Planning road network in mountain forests using GIS and Analytic Hierarchical Process (AHP). *Caspian Journal of environmental sciences*, *8*(2), 151-162.

Mu, E., Pereyra-Rojas, M., (2017). Practical Decision Making an Introduction to the Analytic Hierarchy Process (AHP) Using Super Decisions v2. Springer.

Naghdı, R., Mohammadi Limaei, S. (2009). Optimal Forest Road Density Based on Skidding and Road Construction Costs in Iranian Caspian Forests. Caspian Journal of Environmental Sciences. 7, 2, 79–86.

Pellegrini, M., Grigolato, S. (2013). Spatial multi-criteria decision process to define maintenance priorities of forest road network: an application in the Italian Alpine region. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, *34*(1), 31-42.

Saaty, T. L. (1980). The analytic hierarchy process: planning. *Priority Setting. Resource Allocation, MacGraw-Hill, New York International Book Company*, 287.

Saaty, T.L., (2001). Decision Making with Dependence and Feedback: The Analytic Network Process, 2nd ed. PRWS Publications, Pittsburgh PA.

Saaty, T.L., (2012). Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World. Third, Revised ed. RWS Publications, Pittsburgh.

Sivrikaya, F., Özcan, G. E., Enez, K. (2023). Predicting the susceptibility to *Pityokteines curvidens* using GIS with analytical hierarchy process and, maximum entropy models in fir forests. In Analytic Hierarchy Process-Models, Methods, Concepts, and Applications. Edited by Fabio De Felice and Antonella Petrillo Chapter 8, IntechOpen. pp 153-168.

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.

Thammaboribal, P., Triapthti, N. K., Lipiloet, S. (2025). Using of analytical hierarchy process (AHP) in disaster management: A review of flooding and landslide susceptibility mapping. *International Journal of Geoinformatics*, *21*(4), 177-196.

Unver, S., Ergenc, I. (2021). Safety risk identification and prioritize of forest logging activities using analytic hierarchy process (AHP). *Alexandria Engineering Journal*, *60*(1), 1591-1599.

Vaidya, O. S., Kumar, S. (2006). Analytic hierarchy process: An overview of applications. *European Journal of operational research*, *169*(1), 1-29.

Chapter 11

Stand Basal Area Estimation using Sentinel-2 Time-Series via Diverse Machine Learning Techniques

Semih KUTER[1]*, Alkan GÜNLÜ[2]

Introduction

Forest resources play a crucial role in meeting societal needs in the long term and ensuring the continuous availability of ecological services. Sustainable forest ecosystem management requires detailed information on the fundamental stand characteristics and dynamics. Among these characteristics, the basal area (BA) provides important insights into the structure, condition, and function of forests, and in this respect, it is a priority parameter for sustainable forest management (Appiah Mensah *et al.*, 2020; Bhattarai *et al.*, 2022).

Traditionally, forest inventory data have been collected through field surveys (Hyyppä *et al.*, 2000). When combined with ground-based measurements, inventory data can be determined with the highest accuracy (Gebreslasie *et al.*, 2010; Lu, 2006). However, collecting inventory data through ground measurements across extensive forested regions is a labor-intensive and expensive process (Defibaugh y Chávez and Tullis, 2013; Fatehi *et al.*, 2015). Consequently, remote sensing technologies have emerged as an effective means for estimating stand characteristics across extensive forested areas (Stojanova *et al.*, 2010).

Satellite-based remote sensing, particularly using Landsat and Sentinel-2 (S2) mission data, has become a fundamental approach for estimating stand parameters such as BA, above-ground biomass, number of trees, and stand volume (Aksoy, 2024; Bulut, 2023; Bulut *et al.*, 2023). Landsat's long-term data archive, extending back to the 1970s, enables consistent monitoring of forest changes, recovery, and disturbance over time (Cohen *et al.*, 2010; Kennedy *et al.*,

[1] Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Çankırı, Türkiye, Orcid: 0000-0002-4760-3816, semihkuter@karatekin.edu.tr
[2] Çankırı Karatekin University, Faculty of Forestry, Department of Forestry Engineering, Çankırı, Türkiye, Orcid: 0000-0003-4759-3125, alkangunlu@karatekin.edu.tr

2010). S2, with its high spatial (10–20 m) and spectral resolution, offers enhanced capabilities for discriminating forest species and structural variability (Grabska *et al.*, 2019). As a result, the integration of Landsat and S2 data has become a key methodological framework for modern forest inventory, ecosystem monitoring, and sustainable resource management.

Recent studies using Landsat and S2 imagery have demonstrated that both single-date and multi-temporal datasets can effectively estimate stand parameters such as BA, volume, and above-ground biomass (Forkuor *et al.*, 2020; Ma *et al.*, 2023; Nguyen *et al.*, 2020; Sakici and Gunlu, 2018). Single-date images often yield satisfactory results in structurally stable forests and offer operational simplicity (Forkuor *et al.*, 2020). However, multi-temporal analyses generally provide higher model accuracy and robustness by capturing seasonal and phenological variability, reducing scene-specific noise, and improving feature separability (Grabska *et al.*, 2019). Consequently, while single-date data remains efficient for rapid assessments, multi-temporal approaches are preferred for precise and temporally consistent estimation of stand attributes.

Traditional approaches have relied heavily on Multiple Linear Regression (MLR) models that relate spectral indices or reflectance values to field-measured stand variables. While these regression techniques offer simplicity and interpretability, their linear assumptions often limit performance in complex forest environments characterized by non-linear relationships among spectral, structural, and environmental variables (Foody and Cutler, 2006). To overcome these limitations, recent research has increasingly utilized machine learning (ML) algorithms, which are capable of capturing non-linear and interactive effects in remote sensing data. Notably, Random Forest (RF), Support Vector Regression (SVR), Artificial Neural Networks (ANN), and Gradient Boosting Machines (GBM) have gained the most widespread adoption. (Ali *et al.*, 2015; Belgiu and Drăguţ, 2016; Chen *et al.*, 2023; Maxwell *et al.*, 2018; Zhang *et al.*, 2020).

Despite the growing number of studies utilizing machine learning and remote sensing for forest structure estimation, their performance can vary significantly depending on forest type, stand structure, and ecological conditions. In this context, evaluating and comparing different machine learning algorithms in distinct forest ecosystems is essential to identify the most suitable modeling strategy. The present study focuses on pure oak stands within the Bahçe Forest Planning Unit to investigate the potential of S2 time-series data for estimating BA. By comparing traditional regression and multiple machine learning approaches under identical conditions, this research aims to provide insights into the effectiveness of these methods for operational forest inventory and to

contribute to the development of reliable, cost-efficient, and scalable models for sustainable forest management.

Material and Method

Study Area

The study was conducted within the Bahçe Forest Planning Unit, which is administered by the Osmaniye Forest Enterprise under the Adana Regional Directorate of Forestry, located in southern Türkiye. The planning unit lies between latitudes 37°09′09″–37°17′27″ N and longitudes 36°26′07″–36°41′48″ E (relative to Greenwich). The terrain is characterized by rugged topography and variable elevations that create diverse microclimatic conditions supporting different forest types.

The dominant tree species in the region include *Pinus brutia* Ten., *Pinus nigra* subsp. *pallasiana* (Lamb.) Holmboe, *Cedrus libani* A. Rich., *Juniperus excelsa*, *Pinus pinea*, *Fagus orientalis* Lipsky., *Carpinus betulus*, *Ostrya carpinifolia* Scop., *Platanus orientalis*, *Juglans regia* L., *Quercus petraea* (Matt.) Liebl., and *Quercus cerris* L., accompanied by typical maquis vegetation. The average annual temperature is 18.7°C, while the mean annual precipitation is 614.7 mm, characteristic of a typical Mediterranean climate, which is marked by mild, wet winters and hot, dry summers (Anonymous, 2026; Bayer Altın and Barak, 2017; Gülci *et al.*, 2017).

The study specifically focuses on pure oak stands within the planning unit, where field-based forest inventory measurements were carried out to determine BA values. These measurements were complemented by S2 satellite imagery acquired during the vegetation growth period (April–September) to analyze spectral–structural relationships. A general map showing the location of the study area is presented in Figure 1.

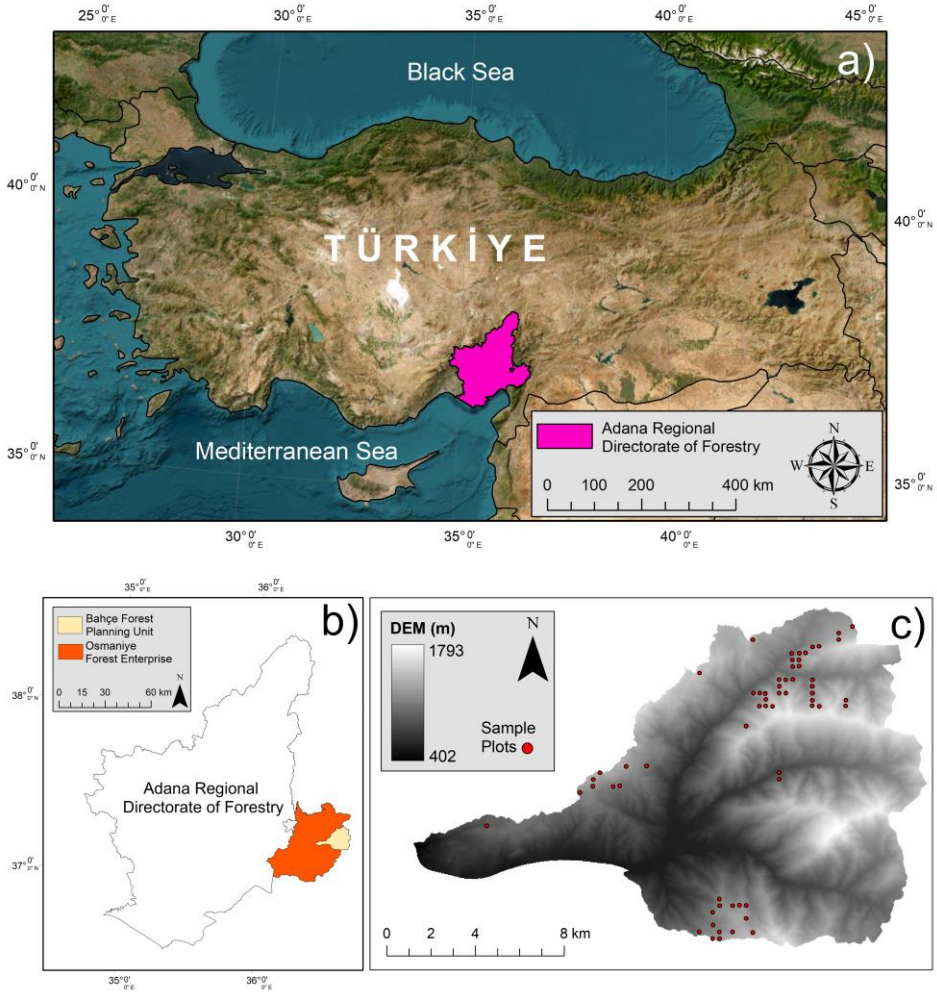

Figure 1. a) Location of the study area: The Adana Regional Directorate of Forestry, b) The Bahçe Forest Planning Unit inside the Osmaniye Forest Enterprise, and c) sample plots over the DEM of the Bahçe Forest Planning Unit.

Ground Data

A total of 61 sample plots were established in pure oak (Quercus *cerris* L.) stands during the 2024 field campaign. For each plot, the diameter at breast height (dbh) of all trees was measured, and the BA of individual trees (m^2) was computed using Equation 1. The sum of individual BA provided the total BA for each plot (m^2). Based on the corresponding plot size, the BA per hectare (m^2 ha^{-1}) was then derived to ensure comparability across plots of different dimensions.

$$BA = \frac{\pi}{4} d_{1.3}^2 \quad (1)$$

In the equation, BA represents the basal area, and $d_{1.3}$ denotes the diameter at breast height (cm).

Remote Sensing Data

Within the scope of the study, S2 satellite images acquired on April 19, May 19, June 18, July 23, August 22, and September 16 of 2025 were utilized. The analysis employed spectral bands (2, 3, 4, and 8) and 20 m (5, 6, 7, 8A, 11, and 12). Surface reflectance images for each band were generated using QGIS 3.40.7. Subsequently, for each band and acquisition date, the mean reflectance values corresponding to the sample plots were extracted through Geographic Information System (GIS) analyses.

Topographic Data

Using the Digital Elevation Model (DEM) of the study area, elevation values for the sample plots were obtained. By employing GIS software, aspect and slope maps of the study area were generated from the DEM data. Subsequently, from these maps, aspect (degree) and slope (%) values corresponding to each sample plot were extracted.

Machine Learning Algorithms

To estimate the BA from the reflectance values of S2 bands and topographic variables, a combination of traditional regression and machine learning algorithms was applied. Each method represents a distinct modeling approach, enabling a comprehensive evaluation of their predictive capabilities and generalization behavior. The selected models include MLR, RF, SVR, ANN, and GBM.

All models were trained and validated using an 80%-20% data split created through simple random sampling, with 80% of the observations randomly selected for training and the remaining 20% allocated for independent testing. This approach ensures an unbiased and consistent evaluation of model performance under identical input conditions.

Multiple Linear Regression

MLR is a commonly used parametric technique for modeling the relationship between a continuous dependent variable and several independent predictors

(Olive, 2017). It assumes a linear relationship between the predictors and the response, expressed in the following general form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon, \quad (2)$$

where y denotes the response variable (i.e., BA), x_i are the predictor variables derived from S2 data, β_I are the regression coefficients representing the contribution of each predictor, and ε is the stochastic error term assumed to follow a normal distribution with zero mean and constant variance (Jobson, 1991).

The model parameters are determined through the ordinary least squares (OLS) approach, which seeks to minimize the sum of the squared deviations between the observed and predicted values. This process provides unbiased estimates under the assumptions of linearity, independence, homoscedasticity, and absence of multicollinearity among predictors. Violation of these assumptions can affect model stability and prediction reliability; therefore, diagnostic analyses such as the variance inflation factor (VIF) are often employed to identify redundant or highly correlated predictors.

In this study, MLR served as a benchmark model to evaluate the relative improvement gained by more flexible machine learning algorithms. While MLR captures only additive and linear effects, it offers clear interpretability through directly estimated coefficients that indicate both the direction and magnitude of predictor influence on BA. This interpretability makes MLR particularly valuable for understanding the fundamental structure of the predictor–response relationships and for assessing the physical plausibility of data-driven patterns derived from more complex models.

Random Forest

RF is an ensemble learning algorithm that does not rely on parameters, aggregating the predictions of several decision trees to enhance model accuracy and generalization. It was introduced by Breiman (2001) as an extension of the bagging (bootstrap aggregation) concept, designed to reduce variance and prevent overfitting in decision tree models. Each tree in the forest is trained on a random bootstrap sample of the original dataset, while a random subset of predictor variables is considered at each node split. This double-randomization strategy ensures that individual trees remain diverse, reducing their correlation and thus increasing the stability of the overall ensemble.

In regression problems, each decision tree provides a numerical prediction, and the RF model outputs the average of all tree predictions. The use of averaging

mitigates the sensitivity of individual trees to noise or extreme values, resulting in smoother and more reliable predictions. The algorithm can model nonlinear and complex interactions among predictors without requiring any assumptions about the underlying data distribution or linearity. Because of these properties, RF has become one of the most popular machine learning techniques in environmental and remote sensing applications, where datasets often exhibit multicollinearity, heterogeneity, and nonlinear behavior.

An additional advantage of RF is its ability to quantify variable importance, typically measured through metrics such as the mean decrease in impurity or the permutation importance (Cutler *et al.*, 2012). These metrics indicate how strongly each predictor contributes to reducing prediction error, helping to identify the most influential spectral, index-based, or textural variables derived from S2 imagery. Moreover, since each tree is trained on a random subset of data, RF inherently provides an internal error estimate—known as the out-of-bag (OOB) error—which serves as an unbiased approximation of model performance without the need for an explicit validation set.

In this study, the RF model was trained on 80% of the available data using simple random sampling, and hyperparameters such as the number of trees, maximum number of splits, and minimum leaf size were optimized to balance accuracy and computational efficiency. The remaining 20% of the data were used for independent testing. The RF approach was expected to provide high predictive accuracy and robustness to noise while maintaining a reasonable level of interpretability through variable importance analysis.

Support Vector Regression

SVR is a supervised learning algorithm derived from the Support Vector Machine (SVM) concept of Vapnik *et al.* (1997), originally developed for classification tasks. It extends the concept of maximizing the margin between data points and separating hyperplane to regression problems, where the objective is to approximate a continuous function that maps predictor variables to a target response. SVR seeks to find an optimal regression function that balances model complexity and prediction accuracy, guided by the principle of structural risk minimization (Smola and Schölkopf, 2004).

Unlike classical regression models that minimize squared errors for all observations, SVR employs an ε-insensitive loss function that ignores small deviations within a specified tolerance (ε) from the regression line. Only the data points lying outside this margin—known as support vectors—influence the model's parameters. This leads to a sparse and computationally efficient model representation that generalizes well, especially when the number of samples is

limited. Owing to this property, SVR has proven to be effective in applications with small to moderately sized datasets, where overfitting can easily occur in more complex models.

To model nonlinear relationships, SVR uses kernel functions that project the predictors into a higher-dimensional feature space, allowing nonlinear structures to be captured through a linear formulation. The radial basis function (RBF) kernel is the most commonly used one due to its ability to represent localized variations in the input–output relationship. The RBF kernel introduces a parameter γ that defines the influence range of individual training samples, thereby controlling the smoothness of the fitted surface.

The overall performance of SVR is determined by three main hyperparameters: the regularization constant (C), the kernel width (γ), and the insensitivity parameter (ε). The parameter C controls the trade-off between minimizing training errors and maintaining model generalization, while ε defines the margin of tolerance and γ adjusts the nonlinearity of the kernel transformation. Careful tuning of these hyperparameters is crucial for achieving an optimal balance between model flexibility and robustness.

In this study, SVR with RBF kernel was trained using 80% of the dataset selected through simple random sampling, with hyperparameter optimization performed on the training set using internal cross-validation. The remaining 20% of the data were used for independent testing. Given its theoretical foundation and ability to extract meaningful patterns from limited data, SVR is particularly well suited for estimating BA from S2 spectral, index-based, and textural predictors.

Artificial Neural Networks

ANNs are flexible, data-driven models inspired by the architecture of biological neurons and designed to learn complex, nonlinear relationships between predictors and response variables (Hornik *et al.*, 1989; Lippmann, 1987). In their basic structure, artificial neural networks (ANNs) consist of a network of interconnected neurons, arranged in an input layer, one or more hidden layers, and an output layer. Each neuron connection is assigned to a numerical weight, which governs the intensity and direction of the influence between them. The model adjusts these weights through multiple iterations during training to minimize the error between the actual and predicted values.

In this study, a feedforward Multilayer Perceptron (MLP) architecture was employed, where information flows unidirectionally from the input layer to the output layer without feedback connections (Hornik *et al.*, 1989). The hidden layers introduce nonlinear transformations through activation functions, enabling

the model to capture complex relationships that are difficult to represent with conventional linear models. Common activation functions contain the rectified linear unit (ReLU), sigmoid, and hyperbolic tangent, which introduce nonlinearity and enhance the model's learning capacity.

Training of the MLP was performed using the backpropagation algorithm (Rojas, 1996), which iteratively updates the connection weights by propagating the error gradient from the output layer back through the network. To avoid overfitting and ensure model generalization, the initial 80% training data—obtained through simple random sampling—was further divided into training and validation subsets. The validation subset was used to monitor performance during training, and the learning process was stopped early once the validation error began to increase, a technique known as early stopping. This procedure prevents the network from memorizing the training data and helps maintain stable performance on unseen observations.

Hyperparameter optimization was carried out on the training set to identify the most suitable network configuration. The tuning process included the number of hidden layers and neurons, the learning rate, the batch size, and the strength of L2 regularization. The hidden layers used the hyperbolic tangent (tanh) activation function, as it provides smooth and bounded nonlinear transformations that facilitate convergence and maintain stable gradients during backpropagation. This combination of controlled network complexity, regularization, and early stopping ensured an optimal balance between fitting accuracy and generalization to unseen data.

Due to its high flexibility and capability to model nonlinear and multivariate dependencies, the ANN approach is well suited for estimating BA from S2 spectral, index-based, and textural predictors, where relationships are inherently complex and not easily captured by traditional regression techniques.

Gradient Boosting Machine

GBM, proposed by Friedman (2001), is an ensemble learning algorithm that builds predictive models in a stage-wise manner by combining a large number of weak learners, typically shallow decision trees. Unlike bagging-based methods such as RF, which train trees independently on bootstrap samples, GBM constructs trees sequentially, where each new tree is fitted to the residual errors of the previous ensemble. This repetitive process enables the model to continuously refine its errors and reduce the total prediction error.

The underlying principle of GBM is gradient descent optimization in function space. In each iteration, the algorithm fits a new tree to the negative gradient of a chosen differentiable loss function (e.g., mean squared error for regression). The

contribution of each tree is scaled by a small learning rate, which controls the step size in the gradient direction and helps prevent overfitting. Although individual trees are weak predictors, their cumulative combination forms a powerful model capable of capturing complex nonlinear relationships (Natekin and Knoll, 2013).

The performance of GBM depends on several key hyperparameters, including the number of boosting iterations (trees), the learning rate, and the maximum depth or number of splits for each tree (Ayyadevara, 2018). Additional parameters, such as the minimum number of observations in a leaf node and the subsampling ratio of data or predictors, can further regulate model complexity and improve generalization. The trade-off between variance and bias is typically managed by using many shallow trees with a small learning rate.

In this study, GBM was trained using 80% of the data obtained through simple random sampling, while the remaining 20% were reserved for independent testing. The training process included hyperparameter optimization on the training subset to identify the most suitable configuration of learning rate, number of trees, and tree depth. The optimization aimed to balance predictive accuracy and computational efficiency while avoiding overfitting.

GBM is particularly well suited for remote sensing applications involving complex and nonlinear predictor–response relationships. Its ability to iteratively focus on the most difficult-to-predict samples and to flexibly model high-order interactions makes it a strong alternative to both RF and ANNs for estimating BA from S2 spectral, index-based, and textural variables.

The Employed Accuracy Metrics for the Assessment of the Models

To evaluate the predictive performance of the developed models, a set of widely used statistical metrics was employed. These metrics provide complementary perspectives on model accuracy, error magnitude, and systematic deviation from observed values. The coefficient of determination (R^2) gauges the proportion of variability in the observed data that is accounted for by the model, reflecting the overall goodness of fit. The root-mean-square error (RMSE) assesses the average size of the prediction errors, with greater emphasis on larger discrepancies, whereas the mean absolute error (MAE) calculates the average absolute difference between predicted and observed values, providing a more balanced perspective on the model's general accuracy. The Bias metric reveals the direction and extent of systematic over- or underestimation by the model. Together, these metrics enable a comprehensive evaluation of predictive performance, facilitating fair comparison among different modeling approaches. The associated formulas for the above-mentioned metrics are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \quad (3)$$

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

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \quad (5)$$

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i), \quad (6)$$

where y_i and $\hat{y}_i$ denote the observed and predicted values, respectively; $\bar{y}$ is the mean of observed values; and n is the total number of samples.

Variable Importance Analysis

In addition to evaluating model accuracy, the relative contribution of each predictor variable to BA estimation was examined. For the RF model, predictor importance was directly obtained from the ensemble's split-based importance scores, which quantify how much each variable contributes to reducing prediction error across decision trees (Boulesteix *et al.*, 2012). For the remaining algorithms —MLR, SVR, ANN, and GBM — variable importance was quantified through a permutation-based approach. In this method, the prediction error was recalculated after randomly shuffling each predictor while keeping others unchanged, and the resulting increase in error was used as an indicator of that variable's importance (Altmann *et al.*, 2010). This unified framework allowed a consistent comparison of predictor influence across all modeling techniques, highlighting the spectral and temporal features most relevant to BA estimation from S2.

Permutation-based variable importance has the advantage of being model-agnostic, as it relies solely on the relationship between predictor perturbation and prediction degradation rather than on model-specific parameters. By randomly permuting the values of a given predictor, its contribution to the model's predictive performance is disrupted while the others remain unchanged. The corresponding increase in prediction error thus reflects the importance of that

variable in maintaining model accuracy (Debeer and Strobl, 2020). This approach ensures comparability across diverse algorithms such as MLR, SVR, ANN, and GBM, providing a consistent and interpretable measure of predictor relevance independent of model complexity or internal structure.

To facilitate interpretability, the variable importance values obtained from each model were normalized and presented as relative scores. The normalized importance represents the proportional contribution of each predictor to the overall model performance, summing to unity within each algorithm. For clarity, only the ten most influential predictors were reported for each model. This ranking highlights the spectral bands and derived indices with the greatest predictive power for BA estimation, allowing a direct comparison of the most informative S2 features across the different regression frameworks. Normalization was achieved by dividing each variable's raw importance value by the sum of all importance scores within the model. This approach transforms absolute values into dimensionless ratios, enabling the identification of predictors that contribute most strongly to model performance regardless of the scale or range of their original importance measures. As a result, the normalized scores provide a consistent basis for comparing predictor relevance across models of differing complexity and structure.

Results and Discussion

The observed and predicted BA values obtained from the test data are presented in Figure 2 for each modelling approach. In addition, information regarding the ten most important variables included in the models and their contributions to model performance is provided in Table 1.

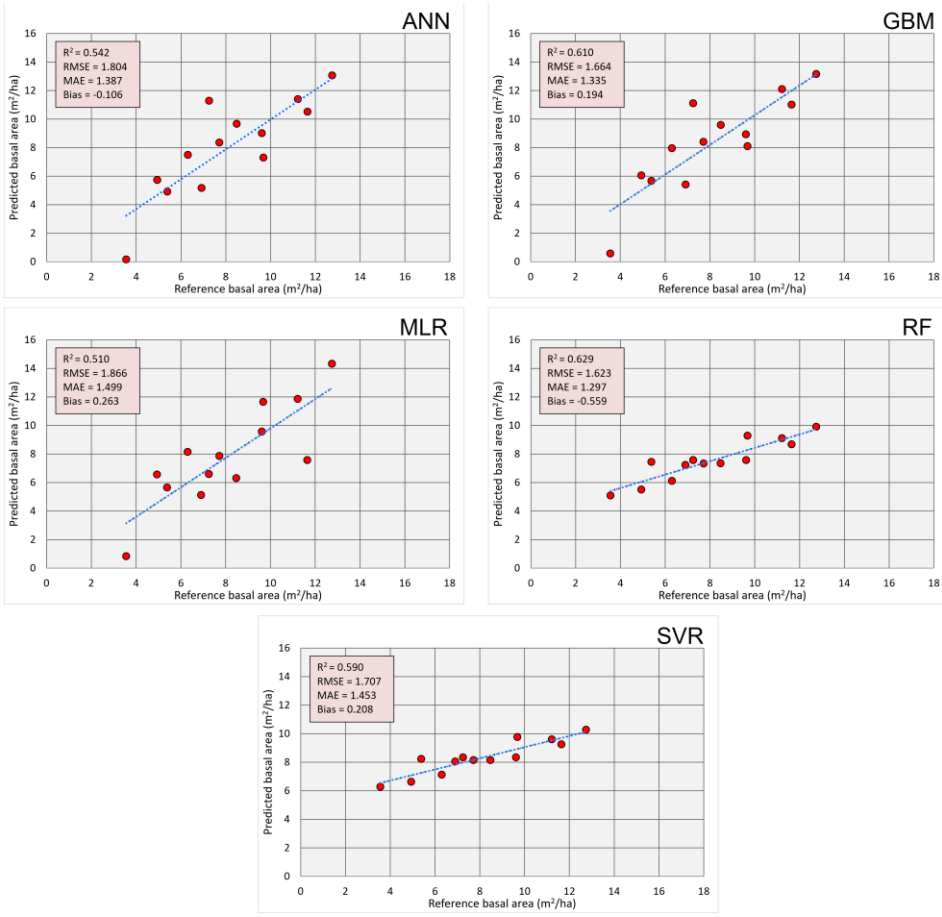

Figure 2. The results on the test dataset for each selected algorithm.

As shown in Figure 2, the RF modeling technique demonstrated the highest performance among the methods employed in the study. This was followed by the GBM modeling technique. The least successful modeling technique was found to be the MLR.

As shown in Table 1, the RF modeling technique, identified as the most successful model, determined the B3 of September S2 image as the most important variable. Similarly, in the MLR and SVR modeling techniques, the most influential variables were different bands of S2 images corresponding to various months. In contrast, for the ANN and GBM modeling techniques, the most important variables were topographic factors, specifically elevation and slope for ANN and GBM, respectively.

Table 1. Normalized importance scores of the top ten predictors for each regression model, indicating their relative contribution to BA estimation from S2.

ANN		GBM		MLR		RF		SVR	
Predictor	**Score**	**Predictor**	**Score**	**Predictor**	**Score**	**Predictor**	**Score**	**Predictor**	**Score**
Elevation	0.187	Slope	0.175	B8_Jul	0.126	B3_Sep	0.104	B2_Aug	0.188
B4_Sep	0.079	B8_Aug	0.148	B6_Dec	0.124	B6_Apr	0.103	B11_Dec	0.145
B5_Apr	0.060	B6_Jun	0.094	B8_Sep	0.100	B3_Dec	0.073	B3_Dec	0.134
B2_Dec	0.054	B3_Apr	0.089	B6_Jun	0.091	B2_Dec	0.055	B5_Apr	0.105
B6_Dec	0.047	B5_Jul	0.088	B7_Sep	0.087	B11_Dec	0.053	B6_Apr	0.081
B8_Dec	0.044	B12_Dec	0.075	B8A_Dec	0.066	B4_Dec	0.043	B12_Dec	0.069
B5_Jun	0.042	B3_May	0.052	B8_Dec	0.061	B7_Apr	0.039	B8_Jul	0.068
B11_Jul	0.041	B8_May	0.048	B7_Dec	0.048	B5_Dec	0.038	B6_Jul	0.058
B8A_Dec	0.038	B2_Dec	0.039	B6_Sep	0.042	Aspect	0.031	B4_Dec	0.058
B6_Apr	0.035	B12_Apr	0.038	B5_Dec	0.032	B12_Dec	0.023	NBANT8	0.032

Several previous studies in the literature have explored the relationships between stand BA and satellite-derived spectral information. Bulut *et al.* (2023) analyzed the relationship between BA values derived from 180 sample plots in pure *Pinus nigra* stands and spectral reflectance data from Landsat 8 (L8) and S2 imagery using and MLR-based approach. The results showed the $R^2 = 0.34$ for L8 and $R^2 = 0.45$ for S2. Similarly, Aksoy (2023) investigated pure *Pinus sylvestris* L. stands and reported $R^2 = 0.48$ for L8 and $R^2 = 0.45$ for S2 using MLR modeling. Demir and Günlü (2024) examined pure poplar (*Populus tremula*) stands and modeled the relationships between S2 reflectance values and BA using MLR, obtaining an $R^2 = 0.342$.

In another study, Aksoy (2024) employed SVM, MLR, Decision Tree (DT), and RF algorithms to model the relationship between BA values derived from 250 sample plots and spectral reflectance from L8 and S2 imagery. The MLR model achieved an R^2 value of 0.24 for both satellite datasets. When topographic variables were incorporated, model performance improved. Using the same modeling techniques, Aksoy (2024) also investigated the relationships between BA and reflectance-based vegetation indices from L8 and S2 data. The reported results were $R^2 = 0.62$ (training) and $R^2 = 0.54$ (testing) for MLR; $R^2 = 0.44$ (training) and $R^2 = 0.46$ (testing) for SVM; $R^2 = 0.64$ (training) and $R^2 = 0.62$ (testing) for DT; and $R^2 = 0.70$ (training) and $R^2 = 0.68$ (testing) for RF.

Aksoy and Günlü (2025) investigated pure *Pinus sylvestris* stands, modeling the relationship between aboveground biomass (AGB) derived from 184 sample plots and spectral reflectance from L8 and S2 imagery using MLR. The reported model performances were $R^2 = 0.52$ for L8 and $R^2 = 0.47$ for S2. In the same study, an ANN model was developed using the predictor set of the best-performing MLR model, where S2 texture metrics served as independent

variables. The models were evaluated separately for training and testing datasets. For MLR, R^2 values were 0.81 (training) and 0.84 (testing), while for ANN they were 0.82 and 0.75, respectively.

Similarly, Vafaei et al. (2018) modeled AGB estimated from 149 sample plots using S2 derived variables and applied RF, SVR, and ANN algorithms. Their findings showed that R^2 values for training and testing were 0.95 and 0.62 for RF, 0.74 and 0.62 for SVR, and 0.72 and 0.58 for ANN, respectively.

In this context, the number of available training samples plays a crucial role in the performance and generalization capability of machine learning models (Fassnacht *et al.*, 2014; Rajput *et al.*, 2023). Algorithms such as RF, GBM, and ANN typically require larger datasets to effectively capture complex, non-linear relationships between spectral predictors and biophysical variables. Conversely, simpler models such as MLR or SVR may achieve reasonable accuracy with smaller sample sizes but are more limited in representing intricate patterns. Therefore, the quantity and diversity of training data directly influence model stability and predictive accuracy, particularly in remote sensing-based forest parameter estimation where field data acquisition is often constrained by cost and accessibility.

Conclusion

In this study, the relationships between BA values computed from field plots and the spectral reflectance values of S2 images acquired during the growing season (April–September) were modeled using ANN, GBM, MLR, SVM, and RF. Evaluation of the test results indicated that the RF model achieved the best performance ($R^2 = 0.610$, RMSE = 1.664 $m^2\ ha^{-1}$). Future studies may improve model accuracy by increasing the number of field plots and incorporating additional predictor variables derived from Landsat and S2 imagery, such as vegetation indices and texture metrics.

Acknowledgements

The authors gratefully acknowledge the General Directorate of Forestry of Turkey for supplying the field data used in this research.

References

Aksoy, H. (2023). *Sinop Orman Bölge Müdürlüğü saf sarıçam meşcerelerinde farklı uzaktan algılama verileri kullanılarak bazı meşcere parametrelerinin modellenmesi.* Çankırı Karatekin Üniversitesi, Doktora Tezi. Çankırı Karatekin Üniversitesi.

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,* 28(8), pp. 2687-2704.

Aksoy, H. and Günlü, A. (2025). UAV and satellite-based prediction of aboveground biomass in scots pine stands: a comparative analysis of regression and neural network approaches. *Earth Science Informatics*, *18*(1), 66.

Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M. and Notarnicola, C. (2015). Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sensing, 7*(12), 16398-16421. doi:10.3390/rs71215841

Altmann, A., Toloşi, L., Sander, O. and Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics,* 26(10), pp. 1340-1347.

Anonymous. (2026). Adana Orman Bölge Müdürlüğü, Osmaniye Orman İşletme Müdürlüğü, Bahçe Orman İşletme Şefliği, Ekosistem Tabanlı Fonksiyonel Orman Amenajman Planı. *Orman Genel Müdürlüğü, Orman İdaresi ve Planlama Dairesi, Ankara.*

Appiah Mensah, A., Petersson, H., Saarela, S., Goude, M. and Holmström, E. (2020). Using heterogeneity indices to adjust basal area – Leaf area index relationship in managed coniferous stands. *Forest Ecology and Management,* 458, pp. 117699.

Ayyadevara, V. K. (2018). Gradient Boosting Machine. In V. K. Ayyadevara (Ed.), *Pro Machine Learning Algorithms : A Hands-On Approach to Implementing Algorithms in Python and R* (pp. 117-134). Berkeley, CA: Apress.

Bayer Altın, T. and Barak, B. (2017). Trends and changes in tropical and summer days at the Adana Sub-Region of the Mediterranean Region, Southern Turkey. *Atmospheric Research,* 196, pp. 182-199.

Belgiu, M. and Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing,* 114, pp. 24-31.

Bhattarai, R., Rahimzadeh-Bajgiran, P., Weiskittel, A., Homayouni, S., Gara, T. W. and Hanavan, R. P. (2022). Estimating species-specific leaf area index and basal area using optical and SAR remote sensing data in Acadian mixed spruce-fir forests, USA. *International Journal of Applied Earth Observation and Geoinformation,* 108, pp. 102727.

Boulesteix, A.-L., Janitza, S., Kruppa, J. and König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *WIREs Data Mining and Knowledge Discovery,* 2(6), pp. 493-507.

Breiman, L. (2001). Random forests. *Machine Learning,* 45(1), pp. 5-32.

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, pp. 101951.

Bulut, S., Günlü, A. and Çakır, G. (2023). Modelling some stand parameters using Landsat 8 OLI and Sentinel-2 satellite images by machine learning techniques: a case study in Türkiye. *Geocarto International,* 38(1), pp. 2158238.

Chen, Y.-N., Fan, K.-C., Chang, Y.-L. and Moriyama, T. (2023). Special Issue Review: Artificial Intelligence and Machine Learning Applications in Remote Sensing. *Remote Sensing, 15*(3). doi:10.3390/rs15030569

Cohen, W. B., Yang, Z. and Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation. *Remote Sensing of Environment,* 114(12), pp. 2911-2924.

Cutler, A., Cutler, D. R. and Stevens, J. R. (2012). Random Forests. In C. Zhang and Y. Ma (Eds.), *Ensemble Machine Learning: Methods and Applications* (pp. 157-175). Boston, MA: Springer US.

Debeer, D. and Strobl, C. (2020). Conditional permutation importance revisited. *BMC Bioinformatics,* 21(1), pp. 307.

Defibaugh y Chávez, J. and Tullis, J. A. (2013). Deciduous Forest Structure Estimated with LIDAR-Optimized Spectral Remote Sensing. *Remote Sensing, 5*(1), 155-182. doi:10.3390/rs5010155

Fassnacht, F. E., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, P. and Koch, B. (2014). Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment,* 154, pp. 102-114.

Fatehi, P., Damm, A., Schaepman, M. E. and Kneubühler, M. (2015). Estimation of Alpine Forest Structural Variables from Imaging Spectrometer Data. *Remote Sensing, 7*(12), 16315-16338. doi:10.3390/rs71215830

Foody, G. M. and Cutler, M. E. J. (2006). Mapping the species richness and composition of tropical forests from remotely sensed data with neural networks. *Ecological Modelling,* 195(1), pp. 37-42.

Forkuor, G., Benewinde Zoungrana, J.-B., Dimobe, K., Ouattara, B., Vadrevu, K. P. and Tondoh, J. E. (2020). Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets - A case study. *Remote Sensing of Environment,* 236, pp. 111496.

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics,* 29(5), pp. 1189-1232.

Gebreslasie, M. T., Ahmed, F. B. and van Aardt, J. A. N. (2010). Predicting forest structural attributes using ancillary data and ASTER satellite data. *International Journal of Applied Earth Observation and Geoinformation,* 12, pp. S23-S26.

Grabska, E., Hostert, P., Pflugmacher, D. and Ostapowicz, K. (2019). Forest Stand Species Mapping Using the Sentinel-2 Time Series. *Remote Sensing, 11*(10). doi:10.3390/rs11101197

Gülci, N., Akay, A. E. and Erdaş, O. (2017). Productivity assessment of alternative timber debarking methods. *Šumarski list,* 141(9-10), pp. 469-476.

Gunlu, A. and Demir, S. (2024). Sentinel-2 uydu görüntüsü kullanılarak saf kavak meşcerelerinde meşcere hacmi ve göğüs yüzeyinin tahmin edilmesi. *Anadolu Orman Araştırmaları Dergisi,* 10(2), pp. 78-86.

Hornik, K., Stinchcombe, M. and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks,* 2(5), pp. 359-366.

Hyyppä, J., Hyyppä, H., Inkinen, M., Engdahl, M., Linko, S. and Zhu, Y.-H. (2000). Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management,* 128(1), pp. 109-120.

Jobson, J. D. (1991). Multiple Linear Regression. In J. D. Jobson (Ed.), *Applied Multivariate Data Analysis: Regression and Experimental Design* (pp. 219-398). New York, NY: Springer New York.

Kennedy, R. E., Yang, Z. and Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sensing of Environment,* 114(12), pp. 2897-2910.

Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine,* 4(2), pp. 4-22.

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

Ma, T., Hu, Y., Wang, J., Beckline, M., Pang, D., Chen, L., Ni, X. and Li, X. (2023). A Novel Vegetation Index Approach Using Sentinel-2 Data and Random Forest Algorithm for Estimating Forest Stock Volume in the Helan Mountains, Ningxia, China. *Remote Sensing, 15*(7). doi:10.3390/rs15071853

Maxwell, A. E., Warner, T. A. and Fang, F. (2018). Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing,* 39(9), pp. 2784-2817.

Natekin, A. and Knoll, A. (2013). Gradient boosting machines, a tutorial. Volume 7 - 2013.

Nguyen, T. H., Jones, S. D., Soto-Berelov, M., Haywood, A. and Hislop, S. (2020). Monitoring aboveground forest biomass dynamics over three decades using Landsat time-series and single-date inventory data. *International Journal of Applied Earth Observation and Geoinformation,* 84, pp. 101952.

Olive, D. J. (2017). Multiple Linear Regression. In D. J. Olive (Ed.), *Linear Regression* (pp. 17-83). Cham: Springer International Publishing.

Rajput, D., Wang, W.-J. and Chen, C.-C. (2023). Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics,* 24(1), pp. 48.

Rojas, R. (1996). The Backpropagation Algorithm. In R. Rojas (Ed.), *Neural Networks: A Systematic Introduction* (pp. 149-182). Berlin, Heidelberg: Springer Berlin Heidelberg.

Sakici, O. and Gunlu, A. (2018). Artificial intelligence applications for predicting some stand attributes using Landsat 8 OLI satellite data: A case study from Turkey. *Applied Ecology and Environmental Research,* 16(4), pp. 5269-5285.

Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing,* 14(3), pp. 199-222.

Stojanova, D., Panov, P., Gjorgjioski, V., Kobler, A. and Džeroski, S. (2010). Estimating vegetation height and canopy cover from remotely sensed data with machine learning. *Ecological Informatics,* 5(4), pp. 256-266.

Vapnik, V., Golowich, S. E. and Smola, A. J. (1997). Support Vector Method for Function Approximation, Regression Estimation and Signal Processing. In M. C. Mozer, M. I. Jordan and T. Petsche (Eds.), *Advances in Neural Information Processing Systems 9* (pp. 281-287): MIT Press.

Vafaei, S., Soosani, J., Adeli, K., Fadaei, H., Naghavi, H., Pham, T. D. and 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.

Zhang, Y., Ma, J., Liang, S., Li, X. and Li, M. (2020). An Evaluation of Eight Machine Learning Regression Algorithms for Forest Aboveground Biomass Estimation from Multiple Satellite Data Products. *Remote Sensing, 12*(24). doi:10.3390/rs12244015

Chapter 12

Forecasting and Optimisation Models in Wood-Based Manufacturing

Hakan AYDOĞAN[1]

Introduction

Wood-based manufacturing, including lumber, wood-based panels (MDF, particleboard, and plywood), pulp and paper, and furniture, forms a core pillar of the bioeconomy. These activities supply indispensable inputs to construction, packaging, and energy systems while contributing to carbon storage and material circularity. Yet firms operate under pronounced market volatility, irregular raw-material availability, and tightening energy and sustainability constraints. Therefore, systematic approaches to forecasting and optimisation are increasingly necessary to sustain cost efficiency, operational resilience, and environmental performance.

Forecasting provides the quantitative basis for anticipating demand, planning production, and detecting recurrent patterns in sectoral output. Among classical time-series methods, the moving average (MA) is widely used to smooth short-run noise and expose longer-run movements, particularly when series exhibit irregularities or temporary shocks typical of wood product markets (Billah et al., 2024; Banaś & Utnik-Banaś, 2021; Kożuch et al., 2023). Moreover, simple exponential smoothing (ES) increases responsiveness by placing higher weight on recent observations, which is useful when market conditions shift with construction cycles, export seasonality, or policy changes. When both trend and seasonality are present, as is frequently the case in panel and furniture production, the Holt–Winters scheme (Hyndman & Athanasopoulos, 2018) offers a natural extension by modelling level, trend, and seasonal components in an internally consistent manner.

For series with evident serial dependence, the autoregressive integrated moving average (ARIMA) framework combines differencing, autoregressive, and moving-average terms in a compact specification. In practice, ARIMA often

[1] Kastamonu University, Faculty of Forestry, Department of Foreset Industry Engineering, Kastamonu, Türkiye, Orcid: 0000-0001-9482-9888, hakanaydogan@kastamonu.edu.tr

performs well for medium-horizon projections of production and export indices under volatility and gradual structural change. Thus, ARIMA provides a transparent benchmark against which more elaborate methods can be assessed. Applications in energy and biomass price analysis and in lumber market forecasting, including studies from Chile, report reliable short-term projections of sawlog prices and export dynamics under these conditions (Mejías Caballero et al., 2023).

While forecasting provides the foresight necessary for planning, operational efficiency depends on how these forecasts are translated into quantitative decisions through optimisation. Classical optimisation methods continue to play a central role in aligning production, inventory, and procurement with forecasted demand. The Economic Order Quantity (EOQ) model, for instance, determines the optimal replenishment size by minimizing the combined costs of ordering and holding inventory. This approach has been successfully applied in small and medium-sized enterprises to enhance stock management and reduce working capital requirements (Hasbullah et al., 2021; Sah & Fıredi-Fulop, 2022). Because demand rarely remains constant in the wood sector, the Reorder Point (ROP) model becomes essential: it identifies the inventory threshold that triggers replenishment while accounting for demand variability and lead time. In stochastic environments, the inclusion of a safety-stock term based on the standard deviation of demand during lead time allows firms to sustain high service levels while preventing costly stockouts.

Beyond inventory control, Linear Programming (LP) techniques have long been used to optimize production schedules, resource allocation, and cutting patterns. LP-based formulations minimize cost or maximize yield subject to process, capacity, and material constraints on throughput (Alotaibi & Nadeem, 2021; Chanda et al., 2022; Kunwar & Sapkota, 2022; Chaudhary & Giri, 2024). In the wood industry, such models have been applied to sawing pattern optimisation, pulp production planning, and multi-plant coordination, demonstrating measurable reductions in waste and improved resource utilization (Hosseini & Peer, 2022; Avilés et al., 2023; Damaševičius & Maskeliūnas, 2025). Recent extensions integrate environmental objectives, incorporating carbon or energy constraints into the optimisation process to align production decisions with sustainability goals (Khahro et al., 2021; Li et al., 2023; Song et al., 2024). This evolution from purely economic to multi-objective optimisation underscores the growing emphasis on balancing profitability with ecological performance in forest-based manufacturing systems.

The combined use of forecasting and optimisation constitutes a comprehensive analytical framework for modern wood-based production

management. Forecasting models such as MA, Exponential Smoothing, Holt–Winters, and ARIMA provide reliable projections of demand and production trends, while optimisation methods including EOQ, ROP, and LP convert these projections into efficient, cost-effective, and sustainable operational strategies. Moreover, their integration enables a closed feedback loop in which forecast errors can be continuously evaluated and decision parameters adjusted accordingly. For example, demand forecasts derived from Holt–Winters or ARIMA models can serve as direct inputs to EOQ or LP formulations, allowing manufacturers to align production volumes and inventory policies with projected market conditions.

Nevertheless, many firms in the wood sector continue to rely on heuristic or experience-based planning due to limited digital infrastructure and data accessibility. Yet empirical studies increasingly indicate that even simple quantitative models, when systematically combined within forecasting–optimisation frameworks, can substantially improve production stability, reduce waste, and enhance responsiveness to market dynamics. Consequently, as digital transformation progresses and high-resolution production data become more available, the integration of classical time-series forecasting with optimisation-based decision frameworks will be central to advancing competitiveness and sustainability in the wood-based manufacturing industry.

Data and Methods

This section outlines the methodological framework used to evaluate forecasting and optimisation techniques for wood-based manufacturing. The chapter employs a synthetic dataset to replicate realistic industrial conditions while preserving analytical flexibility and reproducibility. This design enables controlled tests under varying levels of demand, trend, and seasonality that resemble official statistics such as Eurostat or FAOSTAT, without reliance on proprietary data. We operationalize this approach with a synthetic monthly production index for wood-based panels covering 240 observations from January 2005 to December 2024. The series was engineered to match industrially plausible growth, annual seasonality, and medium-term cycles, with full reproducibility. The methodology is structured around two principal components: (i) forecasting models used to predict production or demand over time, and (ii) optimisation models used to transform forecasts into operational decisions concerning inventory and production.

Data Description

To ensure transparency and replicability, the empirical analysis uses simulated time series that represent production and demand dynamics in the wood-based manufacturing sector. Real world datasets from Eurostat and FAOSTAT inform the choice of magnitudes and periodicities, yet the data analysed here are synthetic and calibrated to emulate the statistical features of industrial series.

Data-generating process: Let $t = 1, \ldots, 240$ index months from 2005-01 to 2024-12. The observed index is built as a superposition of deterministic trend, annual seasonality, a five-year cycle, and idiosyncratic shocks:

$$Y_t = \underbrace{100 + 50\frac{t-1}{239}}_{Trend} + \underbrace{8\sin\left(\frac{2\pi t}{12}\right)}_{Seasonality} + \underbrace{3\sin\left(\frac{2\pi t}{60}\right)}_{Cycle} + \varepsilon_t \quad \varepsilon_t \sim N(0, 2.5^2) \quad (1)$$

The trend rises from 100 to 150 over the sample. The seasonal component has period 12 with amplitude 8. The low-frequency cycle has period 60 months with amplitude 3. Random shocks ε_t capture short-run disturbances. A fixed pseudo-random seed (42) is used for exact reproducibility.

Parameters such as growth rates, seasonal fluctuations, and random shocks were tuned based on empirical findings from earlier forestry time series research (Banaś & Utnik-Banaś, 2021; Heshmatol Vaezin et al., 2022; Damaševičius & Maskeliūnas, 2025).

Before analysis, the simulated series are subjected to the same diagnostic procedures typically used for real industrial data: (i) Stationarity testing via Augmented Dickey–Fuller (ADF) and KPSS tests. (ii) Autocorrelation analysis using ACF and PACF functions to identify potential ARIMA structures. (iii) Outlier and structural break detection through additive outlier (AO) and level shift (LS) tests. All diagnostics are applied to the simulated series to mirror the workflow used for plant or official data. For evaluation, data are split into a training set comprising January 2005–December 2022 (216 months) and a validation set comprising January 2023–December 2024 (24 months) to assess out-of-sample accuracy.

Forecasting Techniques

The forecasting stage employs four established statistical methods widely used for industrial and economic time-series: Moving Average (MA), Exponential Smoothing (ES), Holt–Winters (HW), and Autoregressive Integrated Moving Average (ARIMA). These models were selected for interpretability, methodological diversity, and demonstrated suitability for cyclical production

systems. Seasonality is set to $s = 12$ for monthly data; smoothing and ARIMA orders are selected by minimizing validation RMSE subject to AIC/BIC checks and white-noise residual diagnostics (Ljung–Box), ensuring that remaining autocorrelation is negligible in fitted models.

Moving Average (MA)

The Moving Average model smooths short-term irregularities in simulated demand or production data to reveal the underlying trend component. The n-period moving average is expressed as:

$$MA_t = \frac{Y_{t-1} + Y_{t-2} + \cdots + Y_{t-n}}{n} \tag{2}$$

where Y_t denotes observed production at time t, and n is the number of preceding observations included in the average. MA is particularly suitable for detecting short-run fluctuations in noisy, cyclical datasets similar to those found in real panel production indices (Buongiorno & Zhu, 2015). For multi-step validation, MA(3) yields a constant forecast equal to the last rolling mean, serving as a conservative baseline against which seasonal specifications are assessed.

Exponential Smoothing

Exponential Smoothing enhances adaptability by assigning exponentially decaying weights to older observations:

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t \tag{3}$$

Where $\alpha \in (0,1)$ determines the smoothing intensity. For the simulated series, α is optimised using grid search to minimize RMSE on the validation sample. ES responds quickly to demand surges or shifts often observed in timber and furniture markets (Kożuch et al., 2023; Lamichhane et al., 2023).

Holt-Winters Method

The Holt–Winters model extends exponential smoothing by explicitly modelling level (L_t), trend (T_t), seasonal (S_t) components:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \tag{4}$$

$$T_t = \beta (L_t - L_{t-1}) + (1 - \beta) T_{t-1} \tag{5}$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma) S_{t-s} \tag{6}$$

$$\hat{Y}_{t+h} = (L_t + hT_t) S_{t-s+h} \tag{7}$$

We use the additive specification with $s = 12$ and classical initialization, which captures the strong annual seasonality in wood market statistics.

ARIMA Models

For a flexible stochastic representation, the ARIMA (p, d, q) is specified as:

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)\varepsilon_t \tag{8}$$

Model orders (p, d, q) are selected using Akaike (AIC) and Bayesian (BIC) and residuals are checked for white-noise properties via Ljung–Box. The ARIMA framework captures persistence and short-run autocorrelation commonly observed in manufacturing production indices (Li et al., 2023; Mejías Caballero, 2023).

Forecast accuracy is compared using MAPE, RMSE, and Theil's U-statistic. We also report R_ρ^2 the squared Pearson correlation between actuals and forecasts in the validation window, which lies in $[0,1]$ and complements error metrics by summarizing co-movement. The best-performing method provides simulated forecasts that serve as direct inputs for optimisation analysis.

Optimisation Models

Forecast results are then embedded into optimisation models that guide inventory and production decisions. These models are applied to synthetic demand values derived from the forecasting stage so that simulation and optimisation remain methodologically coherent. All optimisation inputs are taken from the validation-calibrated forecasting layer to ensure internal consistency.

Economic Order Quantity (EOQ)

The EOQ model determines the cost-minimizing order quantity:

$$EOQ = \sqrt{\frac{2DS}{H}} \tag{9}$$

where D represents simulated annual demand, S the ordering cost per replenishment, and H the annual holding cost per unit. For sensitivity analysis we

set $D = 12{,}000$ units/year and $S = 500$ per order, and vary the holding cost $H\ on\ [10{,}60]$ to quantify cost-driven changes in lot size. Parameter values are calibrated based on cost structures documented in wood manufacturing case studies.

Reorder Point (ROP)

The ROP model identifies the point at which new orders should be placed:

$$ROP = dL + Z\sigma_L \tag{10}$$

where d is mean simulated demand per period, L is lead time, Z the service level factor, and σ_L the standard deviation of simulated demand during lead time. We use $L = 2$ months and $Z = 1.65$ with σ_L estimated from the validation data. This stochastic formulation ensures high service reliability even under artificial demand volatility.

Linear Programming (LP)

The LP allocates resources efficiently based on simulated demand forecasts. The objective function is:

$$Z_{max} = \sum_{i=1}^{n} c_i x_i \tag{11}$$

$$subject\ to: \sum_{i=1}^{n} \alpha_{ij} x_i \leq b_j, \quad x_i \geq 0 \tag{12}$$

Where $c_i$are profit coefficients, α_{ij} represents input-output relationships, and b_j are resource limits. This model reproduces the trade-offs between production capacity, cost, and inventory observed in empirical wood-processing plants.

Integrated Simulation: Forecast and Optimisation Framework

The analytical design combines simulated time-series generation, statistical forecasting, and optimisation within a unified experimental framework. Forecasts from MA, ES, Holt–Winters, and ARIMA inform EOQ, ROP, and LP stages, establishing a closed-loop environment in which each methodological layer interacts with the others. This structure supports systematic sensitivity tests on demand volatility, cost ratios, and service-level requirements, enabling robust

comparison of model performance across operational scenarios. Such integrated simulation, forecast and optimisation frameworks are increasingly advocated in industrial analytics to evaluate methodological robustness prior to full-scale implementation with proprietary data.

Results and Discussion

This section presents the analytical outcomes of the simulated experiment integrating forecasting and optimisation models for wood-based manufacturing. The results are interpreted in three stages: (i) evaluation of forecasting model performance, (ii) derivation of optimal decisions from optimisation models, and (iii) synthesis of forecasting–optimisation interactions and managerial implications. All statistical computations are based on the simulated dataset described previously, which mirrors the temporal and structural dynamics of typical industrial production indices.

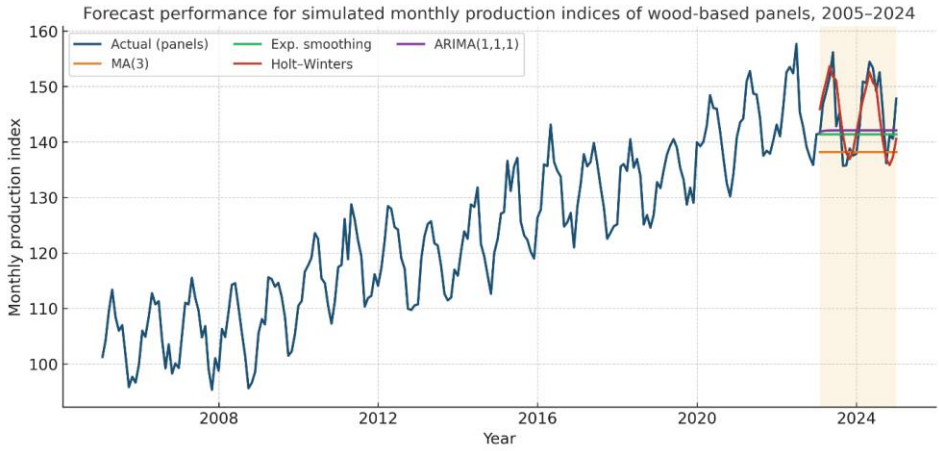

Figure 1. Forecast performance for simulated monthly production indices of wood-based panel, 2005-2024

Figure 1 shows that the Holt–Winters specification reproduces both the level and the seasonal amplitude of the series in the 24-month validation horizon. The hold-out ranking is consistent with visual inspection. Holt–Winters attains MAPE 2.25% and RMSE 3.97. ARIMA(1,1,1) follows with MAPE 3.98% and RMSE 6.99. Exponential smoothing records MAPE 4.10% and RMSE 7.33. The three-month moving average is weakest with MAPE 5.12% and RMSE 9.40. Holt–Winters also preserves the observed seasonal peaks, while ARIMA attenuates oscillations and the level-only methods yield flat paths. The implication is that when seasonality is persistent, a seasonal model should be the default baseline.

To quantify these visual differences and document the size of the accuracy gains that justify model choice, we report the validation metrics in Table 1.

Table 1. Out-of-sample forecast accuracy, validation period 2023–2024

Model	RMSE	MAPE (%)	MAE	Theil's U	R^2 $(corr^2, validation)$
Holt–Winters	3.971	2.248	3.261	0.719	0.71
ARIMA (1,1,1)	6.991	3.984	5.883	1.300	0.52
Exponential Smoothing	7.325	4.099	6.075	1.362	0.38
MA	9.401	5.122	7.650	1.744	0.22

*R^2 column reports squared pearson correlation between actual and forecasts in the validation window.

Table 1 shows three quantitative findings. First, Holt–Winters lowers RMSE by 57.8 percent relative to the moving average, by 45.8 percent relative to exponential smoothing and by 43.2 percent relative to ARIMA. Second, Theil's U equals 0.719 for Holt–Winters, which indicates improvement over a naive benchmark; the other models exceed 1. Third, the squared correlation in validation is 0.71 for Holt–Winters compared with 0.52 for ARIMA, 0.38 for exponential smoothing and 0.22 for the moving average, confirming superior month-to-month tracking.

Having established which forecast is most reliable, we next examine how the forecasted demand scale interacts with cost parameters to determine economically rational order sizes, which motivates the EOQ sensitivity analysis in Figure 2.

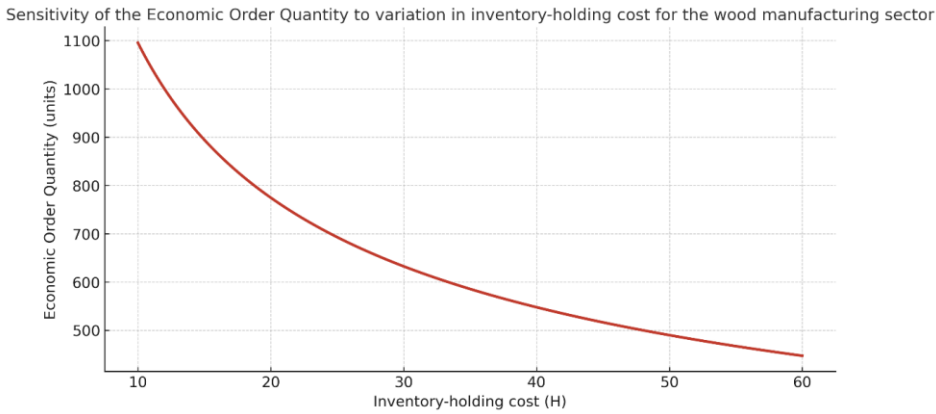

Figure 2. Sensitivity of the EOQ to variation in inventory-holding cost for the wood manufacturing sector

Figure 2 shows the expected monotone and concave relationship $EOQ = \sqrt{\frac{2DS}{H}}$ for $D = 12.000$ units and $S = 500$ per order. At $H = 20$ the optimal lot size is $EOQ \approx 775$ units. Lowering the carrying cost to $H = 15$ increases $EOQ\ to \approx 894$ units. Raising the cost to $H = 40$ reduces EOQ to approximately 548 units, and at $H = 60$ to 447 units. Doubling H from 20 to 40 lowers EOQ by 29.3 percent. The practical message is clear. Accurate measurement of carrying costs is decisive, since EOQ varies with the square root of the inverse of H, while modest forecast error has relatively small influence on the lot size.

Because ordering policy also depends on when to place an order and how to protect service levels against forecast dispersion, we complement EOQ with reorder-point calculations in Table 2.

Table 2. Replenishment policies implied by validation forecasts

Policy	Inputs	Results
EOQ reference	D=12.000, S=500, H=20	$EOQ \approx 775\ units\ per\ order$
ROP deterministic	d=145.16, L=2	ROP = dL = 290.32 units
ROP stochastic	d=145.16, L=2, Z=1.65, $\alpha_L \approx 6.41$	ROP = 300.89 units

Table 2 shows that introducing a 95 percent service target increases the reorder point by 10.57 units relative to the deterministic trigger, which is a rise of 3.6 percent. This is a modest inventory increment that produces a large reduction in stockout risk. EOQ addresses how much to order given costs and scale. ROP translates forecast uncertainty into safety stock during lead time. The two rules are complementary.

With ordering decisions quantified, the final step is to convert monthly demand forecasts into plant-level production plans under capacity and material constraints, which we implement with a linear program summarized in Figure 3.

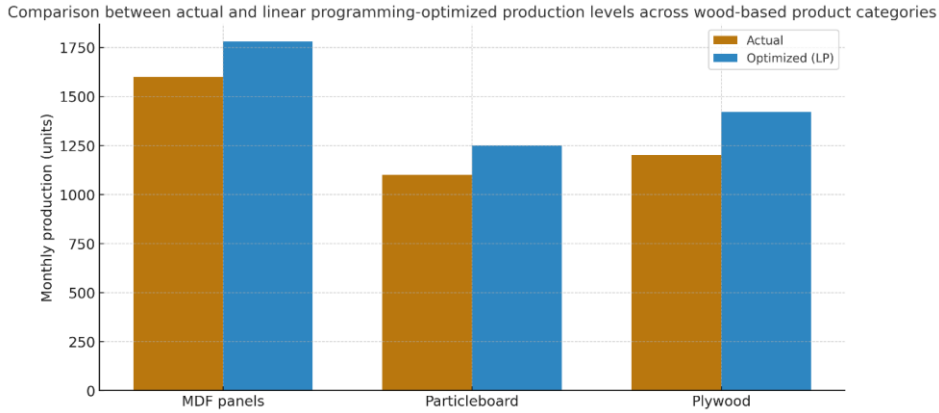

Figure 3. Comparison between actual and LP-optimised production level across wood-based products

Figure 3 shows that the optimisation rebalances capacity in favour of higher-return products. MDF panels rise from 1,600 to 1,780 units per month, which is plus 11.3 percent. Particleboard increases from 1,100 to 1,250 units, which is plus 13.6 percent. Plywood increases from 1,200 to 1,420 units, which is plus 18.3 percent. Aggregate monthly output rises from 3,900 to 4,450 units, an improvement of 14.1 percent. These gains arise from reallocating capacity toward products with superior margin-to-constraint ratios rather than from expanding inventory.

Conclusion

This chapter examined how classical forecasting and optimisation methods perform on a simulated monthly index representative of wood-based panel production. The data were designed to reflect a steady upward trend with pronounced annual seasonality, which mirrors the operating environment of many mills and integrated plants. The analysis combined four forecast models with three optimisation tools in order to build a verifiable framework from prediction to decisions on ordering and production.

The evidence supports three central messages. First, model choice should follow the structure of the data. When seasonality is persistent, a seasonal specification should be the baseline. In our setting, Holt–Winters achieved the strongest out-of-sample accuracy with MAPE near 2.25 percent and RMSE about 3.97. ARIMA(1,1,1) provided a compact comparator but attenuated seasonal amplitude and trailed Holt–Winters on both error metrics and squared correlation in the validation window. Moving average and level-only exponential smoothing

were useful as baselines for robustness, yet they did not reproduce the seasonal pattern and produced nearly flat forward paths.

Second, forecasts must be translated into operational policies. EOQ clarifies how much to order once demand scale and cost parameters are specified. The sensitivity curve confirmed the theoretical relationship EOQ equals the square root of 2DS over H and showed a 29.3 percent reduction in lot size when the holding cost doubled from 20 to 40 currency units. Reorder points determine when to order and are the place where forecast dispersion matters most. Using a 95 percent service target increased the reorder point by 3.6 percent relative to the deterministic trigger. This modest increase in buffer stock is sufficient to materially reduce stockouts during the lead time.

Third, optimisation converts predictive accuracy into value at the plant level. The linear program that allocates production across MDF panels, particleboard, and plywood raised aggregate monthly output from 3,900 to 4,450 units, which corresponds to an improvement of 14.1 percent under the assumed margins and constraints. These gains come from reallocating capacity toward products with superior margin-to-capacity ratios rather than from expanding inventory. The result illustrates the practical role of an explicit optimisation layer that is informed by coherent monthly forecasts.

The study relied on simulated data to ensure transparency and reproducibility. This choice allowed full control over trend, seasonality, and noise, and made it possible to attribute performance differences to model structure rather than to data idiosyncrasies. For applied work with plant data, two extensions are natural. Seasonal ARIMA or seasonal ETS should be tested alongside Holt–Winters when calendar effects, shutdowns, or regime shifts are present. A stochastic or robust production model can represent uncertainty in costs, yields, and energy prices, and can include explicit service level penalties.

There are several practical implications for managers. Benchmark a seasonal forecasting model as a default for panel lines and validate it against a compact ARIMA to check stability. Calibrate EOQ with careful measurement of carrying costs since these parameters dominate the lot-size outcome. Use reorder points that incorporate demand variance during lead time in order to protect service levels at moderate inventory cost. Couple forecasts to a capacity-constrained planning model so that production mix is determined by economic trade-offs rather than by static rules or historical shares.

The workflow documented here establishes a clear trajectory from empirical data to operational decisions. Forecasts provide credible monthly targets, which are then used to set EOQ and ROP policies; these policies are converted into feasible production schedules via a constrained linear program.

The presented approach is straightforward to implement, relies on well-established statistical and optimisation methodologies, and is transparent enough to be audited and refined as richer data become available. Future research should focus on the integration of real-time instrumentation and Manufacturing Execution System (MES) data. In addition, evaluating hybrid approaches that couple seasonal time-series models with pattern-recognition tools for robust outlier handling warrants careful study. Finally, the incorporation of environmental objectives, such as energy consumption and emissions, into multi-objective production planning constitutes a critical area for subsequent work.

References

Alotaibi, A., & Nadeem, F. (2021). A review of applications of linear programming to optimize agricultural solutions. International Journal of Information Engineering and Electronic Business, 10(2), 11.

Avilés, F. N., Etchepare, R. M., Aguayo, M. M., & Valenzuela, M. (2023). A mixed-integer programming model for an integrated production planning problem with preventive maintenance in the pulp and paper industry. Engineering Optimization, 55(8), 1352-1369.

Banaś, J., & Utnik-Banaś, K. (2021). Evaluating a seasonal autoregressive moving average model with an exogenous variable for short-term timber price forecasting. Forest Policy and Economics, 131, 102564.

Billah, M. M., Sultana, A., Bhuiyan, F., & Kaosar, M. G. (2024). Stock price prediction: comparison of different moving average techniques using deep learning model. Neural Computing and Applications, 36(11), 5861-5871.

Chanda, R., Pabalkar, V., & Gupta, S. (2022). A study on application of linear programming on product mix for profit maximization and cost optimization. Indian Journal of Science and Technology, 15(22), 1067-1074.

Chaudhary, S., & Giri, S. R. (2024). Enhancing production efficiency through integer linear programming-based production planning. Journal of Engineering Issues and Solutions, 3(1), 63-75.

Damaševičius, R., & Maskeliūnas, R. (2025). Modeling Forest Regeneration Dynamics: Estimating Regeneration, Growth, and Mortality Rates in Lithuanian Forests. Forests, 16(2), 192.

Damaševičius, R., & Maskeliūnas, R. (2025). Optimising the wood supply chain for enhanced furniture industry efficiency. In Supply Chain Forum: An International Journal (pp. 1-16). Taylor & Francis.

Hasbullah, H., Mustarih, M. M., & Wibowo, A. A. (2021). Improving material shortage for small-medium enterprises (SME) in pest control industry. Journal of Industrial Engineering & Management Research, 2(3), 62-71.

Heshmatol Vaezin, S. M., Moftakhar Juybari, M., Sadeghi, S. M. M., Banaś, J., & Marcu, M. V. (2022). The seasonal fluctuation of timber prices in Hyrcanian temperate forests, northern Iran. Forests, 13(5), 761.

Hosseini, S. M., & Peer, A. (2022). Wood products manufacturing optimization: A survey. IEEE Access, 10, 121653-121683.

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts, Melbourne.

Khahro, S. H., Kumar, D., Siddiqui, F. H., Ali, T. H., Raza, M. S., & Khoso, A. R. (2021). Optimizing energy use, cost and carbon emission through

building information modelling and a sustainability approach: A case-study of a hospital building. Sustainability, 13(7), 3675.

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.

Kunwar, R., & Sapkota, H. P. (2022). An introduction to linear programming problems with some real-life applications. European Journal of Mathematics and Statistics, 3(2), 21-27.

Lamichhane, S., Mei, B., & Siry, J. (2023). Forecasting pine sawtimber stumpage prices: A comparison between a time series hybrid model and an artificial neural network. Forest Policy and Economics, 154, 103028.

Li, C., Liu, S., & Yang, G. (2023). Research on material demand analysis of manufacturing industry based on time series model—ARIMA. Academic Journal of Computing & Information Science, 6(8), 131-137.

Li, Y., Liu, Z., Sang, Y., Hu, J., Li, B., Zhang, X., ... & Zheng, W. (2023). Optimization of integrated energy system for low-carbon community considering the feasibility and application limitation. Applied Energy, 348, 121528.

Mejías Caballero, W., Gysling Caselli, J., Kahler González, C., Soto Aguirre, D., & Pardo, V. (2023). Using ARIMA models to project sawlogs and sawn wood prices in the Chilean construction materials market. World Conference on Timber Engineering 2023.

Sah, G. G., & Furedi-Fulop, J. (2022). The effects of proper inventory management on the profitability of SMEs. Technium Soc. Sci. J., 32, 340.

Song, J., Zhang, Z., Mu, Y., Wang, X., Chen, H., Pan, Q., & Li, Y. (2024). Enhancing envrionmental sustainability via interval optimization for low-carbon economic dispatch in renewable energy power systems: Leveraging the flexible cooperation of wind energy and carbon capture power plants. Journal of Cleaner Production, 442, 140937.