Karadeniz Fen Bilimleri Dergisi, 15(2), 688-701, 2025. DOI: 10.31466/kfbd.1570632

KFBD

Karadeniz Fen Bilimleri Dergisi
The Black Sea Journal of Sciences
ISSN (Online): 2564-7377 https://dergipark.org.tr/tr/pub/kfbd

Araştırma Makalesi / Research Article

Forecasting the Tobacco Market in Türkiye with Artificial Neural Networks

Aysel TOPŞİR[1*], Ferdi GÜLER[2]

Abstract

This study aims to forecast the future dynamics of tobacco policies in Türkiye using artificial neural networks. Tobacco production, area harvested, and yield data from 1961 to 2022 were comprehensively analyzed to understand the complex relationships among these variables. The results indicate that, while tobacco production and harvested area are expected to decline gradually between 2023 and 2027, yield will significantly increase. This trend reflects the positive impact of technological advancements and effective agricultural policies. Time series forecasting was conducted using DeepDenT software. These forecasts provide valuable insights for the sustainability and strategic planning of tobacco farming. In addition to forecasting, the study applied the linear Granger causality test to assess relationships between the variables. However, no statistically significant causality was found, suggesting that tobacco production is influenced by complex, non-linear dynamics. This implies that conventional linear models may be insufficient to capture the true nature of the production process. Overall, the study offers critical insights into long-term trends in tobacco agriculture and contributes to policy development. It supports producers in making informed, strategic decisions and enhances understanding of the sector's sustainability and economic stability. Thus, the study offers a new perspective on optimizing production through data-driven approaches and advanced modeling.

Keywords: Forecasting, Analyzing Tobacco, Artificial Neural Networks.

Türkiye'de Tütün Piyasasının Yapay Sinir Ağları ile Tahmin Edilmesi

Öz

Bu çalışma, Türkiye'deki tütün politikalarının gelecekteki dinamiklerini yapay sinir ağları kullanarak öngörmeyi amaçlamaktadır. 1961–2022 yılları arasındaki tütün üretimi, hasat alanı ve verim verileri, bu değişkenler arasındaki karmaşık ilişkileri anlamak amacıyla kapsamlı bir şekilde analiz edilmiştir. Sonuçlar, 2023 ile 2027 yılları arasında tütün üretimi ve hasat alanının kademeli olarak azalmasının beklendiğini, buna karşılık verimin önemli ölçüde artacağını göstermektedir. Bu eğilim, teknolojik gelişmelerin ve etkili tarım politikalarının olumlu etkilerini yansıtmaktadır. Zaman serisi tahminleri derin dendritik yapay sinir ağları (DeepDenT) yazılımı kullanılarak gerçekleştirilmiştir. Bu tahminler, tütün tarımının sürdürülebilirliği ve stratejik planlaması açısından değerli bilgiler sunmaktadır. Tahminlerin yanı sıra, çalışmada değişkenler arasındaki ilişkileri değerlendirmek amacıyla doğrusal Granger nedensellik testi uygulanmıştır. Ancak, istatistiksel olarak anlamlı bir nedensellik bulunamamıştır; bu durum, tütün üretiminin karmaşık ve doğrusal olmayan dinamiklerden etkilendiğini göstermektedir. Bu da geleneksel doğrusal modellerin üretim sürecinin gerçek doğasını yeterince yansıtamayabileceğini ima etmektedir. Genel olarak, bu çalışma tütün tarımındaki uzun vadeli eğilimlere dair kritik bulgular sunmakta ve politika geliştirme süreçlerine katkı sağlamaktadır. Üreticilerin bilinçli ve stratejik kararlar almalarına destek olmakta ve sektörün sürdürülebilirliği ile ekonomik istikrarına ilişkin anlayışı derinleştirmektedir. Böylece, veri temelli yaklaşımlar ve ileri düzey modelleme teknikleriyle üretim süreçlerinin optimize edilmesine yönelik yeni bir bakış açısı sunmaktadır.

Anahtar Kelimeler: Öngörü, Tütün Analizi, Yapay Sinir Ağları.

[1]Yildiz Technical University, Department of Industrial Engineering, İstanbul, Türkiye, aysel.topsir@std.yildiz.edu.tr
[2]Giresun University, Department of Statistics, Giresun, Türkiye, ferdiiguler@gmail.com

[*]Sorumlu Yazar/Corresponding Author

Geliş/Received: 20.10.2024 **Kabul/Accepted:** 16.05.2025 **Yayın/Published:** 15.06.2025

Topşir, A., Güler, F. (2025). Forecasting the Tobacco Market in Türkiye with Artificial Neural Networks. *Karadeniz Fen Bilimleri Dergisi*, 15(2), 688-701.

1. Introduction

Türkiye holds a significant position in global tobacco production. Tobacco cultivation, especially concentrated in the Aegean and Black Sea regions, makes a substantial contribution to the country's economy and agriculture. In recent years, fluctuations have been observed in the harvested tobacco areas and production volumes. The primary causes of these changes include agricultural policies, global market conditions, and climate change. The efficiency of Türkiye's tobacco production can vary considering these factors, making it increasingly important to forecast future production levels. Tobacco farming holds a crucial place worldwide. For instance, in Türkiye, tobacco production is predominantly based on family labor (Bilir et al., 2010). However, in recent years, the cultivation areas have been shrinking, and younger generations show less interest in this activity (Gumus and Gumus, 2005). Particularly, oriental tobacco produced in Türkiye is recognized for its high quality in global markets and is in demand (Karabacak, 2017). Despite this, Türkiye's position in tobacco production has declined over the past few years (Karabacak, 2017). Although tobacco production in Türkiye has decreased, tobacco consumption remains high in certain regions, contributing to the global problem of the harmful health effects of tobacco use.

While Türkiye was an important tobacco producer and exporter, it began importing tobacco in 1989 (Gumus and Gumus, 2005). The increasing tobacco consumption and changes in cigarette demand have turned Türkiye into a significant tobacco importer (Gumus and Gumus, 2005). Alongside this, rising tobacco use is a leading risk factor for several non-communicable diseases, including cancer, respiratory diseases, and cardiovascular diseases (Asi and Gozum, 2020). Globally, more than 8 million people die each year due to tobacco consumption, with most of these deaths directly attributable to tobacco use (Asi and Gozum, 2020). Tobacco products remain one of the most critical global health issues and are a focal point of health policies in many countries (Pesen et al., 2021). Tobacco uses and production are part of an international problem, with economic, social, and health-related impacts. Therefore, anti-tobacco policies continue to be a critical focus at both national and international levels (Pesen et al., 2021). Estimating future trends in tobacco production is particularly important for producers, as such forecasts play a crucial role in managing supply-demand balance, pricing, and trade policies. With increasing fluctuations in production and global market conditions, advanced forecasting methods have become necessary in recent years. In this context, artificial neural networks have emerged as a powerful tool for handling nonlinear data and providing forecasts for complex production processes.

Machine learning (ML) has recently been effectively utilized across various domains, including bioinformatics (Agraz et al., 2024), scientific ML (Agraz, 2024), energy studies (Magazzino et al., 2021), weather prediction (Shi et al., 2015), and forecasting (Voyant et al., 2017). Time series forecasting problems have been solved using modern methods like artificial neural networks (Alpaslan et al., 2012). Although ANNs offer significant advantages for forecasting problems, determining their architectural structures remains a debated issue (Alpaslan et al., 2012). Incorporating such models into time series forecasting, especially for economic and social contexts such as tobacco consumption predictions, has become quite common in recent years (Erilli et al., 2010). Studies using ANNs for nonlinear predictions have achieved notable success (Erilli et al., 2010).

Over the last few years, deep learning models have come under consideration for their ability to detect complex relationships and long-distance dependencies. Variants of recurrent neural networks (RNNs) (long short-term memory (LSTM) gated recurrent unit (GRU)), convolutional neural networks (CNNs), and transformer models have been suggested for multivariate time series data processing. For both univariate and multivariate, single-step or multi-step time series forecasting (TSF) situations, these models have been shown to improve accuracy (Liu and Wang, 2024). Zhang et al. (2020) used machine learning methods such as backpropagation neural networks (BPNN) and random forests (RF) with hyperspectral imaging to predict tobacco yields and detect diseases. The study concluded that these methods significantly improved the accuracy of yield predictions and allowed for the early detection of diseases. Celik (2020) compared artificial neural networks and the multiplicative decomposition method to model and forecast tobacco production in Türkiye. In this study, years were used as input parameters, while tobacco production quantity was used as the output parameter. The results showed that the ANN model yielded better results with lower mean squared error (MSE) and mean absolute error (MAE) values. It was forecasted that by 2025, tobacco production would increase by 101.53% compared to 2019. Chen et al. (2024) used machine learning models based on spatial data and external variables to predict tobacco yields. In their study, models such as ARIMAX, ANN, and support vector regression (SVR) were compared. The inclusion of spatial data and external variables enhanced the accuracy of yield predictions, and the results demonstrated that these models hold significant potential for the management of tobacco production.

In this study, data on tobacco production, area harvested, and yield between 1961 and 2022 are analyzed to provide forecasts for the next five-year period. These three variables are used to create forecasting models employing artificial neural networks with lagged data. The linear Granger causality test identifies linear relationships between variables in time series data (Sameshima and Baccala, 2016). This test evaluates the extent to which variables are related to each other and whether these relationships are linear in nature, as reflected in model performance (i.e., root mean squared error (RMSE) values). In this study, the linear Granger causality test was applied based on the RMSE error metric provided by the DeepDenT application to determine whether there are linear relationships among the variables production, area harvested and yield. This approach contributes to understanding how the model operates, and which variables may have stronger influences on one another. This approach aims to enhance our understanding of the factors influencing tobacco production processes and to contribute to more evidence-based agricultural policies and more realistic forecasts.

Ultimately, this study will enable a better assessment of changes in production processes, aiding both producers and policymakers in making more strategic decisions. By optimizing agricultural planning and resource use, it is expected that the tobacco sector and agricultural policies will become more resilient to economic uncertainties.

This study includes sections that enhance the understanding of tobacco production dynamics. The first section addresses the reasons behind the fluctuations in tobacco production in Türkiye, along with the general literature on tobacco production and forecasting in tobacco production, as well as the overall objectives of the study. The second section presents the materials and methods, detailing the data sources and techniques used. The third section elaborates on the findings and their significance, while the fourth section discusses the results and their implications in detail. This structure supports the study's objectives regarding the tobacco market in Türkiye and, accordingly, the forecasting targets related to tobacco production, area harvested, and yield.

2. Materials and Methods

2.1. Dataset

This study aims to present trends in the Turkish tobacco market by utilizing data on tobacco production, area harvested, and yield from 1961 to 2022. It seeks to explore the relationship between these three variables and to forecast the expectations for production, area harvested, and yield for the years 2023-2027 through predictive modeling. The data used in this study were obtained from secondary sources provided by reliable institutions such as the Food and Agriculture Organization of the United Nations (FAO) (Anonymous, 2021) and the Turkish Statistical Institute (TurkStat) (Anonymous, 2022). The key subjects of analysis in the dataset are annual tobacco production, the total area planted with tobacco, and yield per hectare in Türkiye. While the data on tobacco production help explain the fluctuations over the years, agricultural policy, climate change, and conditions in the international market are among other factors considered in evaluating these changes. The data on harvested tobacco areas illustrate how much land is used for tobacco production and how this area has changed over time. Tobacco yield represents production per hectare, which is a crucial factor in measuring agricultural yield. Future production trends have been forecasted through time series analyses based on historical data, and the pairwise relationships between these three variables have been examined using artificial neural networks. In the analyses conducted in this study, the automatic forecasting method based on DeepDenT, an open-access MATLAB application of a deep dendritic ANN proposed by Egrioglu and Bas (2024), was used as the automatic forecasting method.

The data presented in Table 1 detail the annual changes in tobacco production, area harvested, and fluctuations in yield. This table allows for an examination of the long-term trends in tobacco production and the interrelationships between these variables. These data were utilized in the forecasting model as training data (1961-2010, 80%), validation data (2011-2016, 10%), and test data (2017-2022, 10%).

Table 1. Historical data on tobacco production, harvested area, and yield in Türkiye (1961–2022).

Years	Production	Area Harvested	Yield
1961	101407	140625	7211
1962	89793	149346	6012
.	.	.	.
2022	82250	75051	10959

2.2. Forecasting

Forecasting is defined as the process of predicting the future values of a variable based on past and present data (Ataseven, 2013). Accurate forecasts play a critical role, especially in decision-making processes, as successful predictions enhance the accuracy of decisions and the efficiency of outcomes. This has led to an increasing interest in forecasting models (Ataseven, 2013). Time series forecasting using artificial neural networks has been widely applied in the literature (Alpaslan et al., 2012).

Time series are the fundamental building blocks of forecasting studies; however, in recent years, more flexible and effective methods have made significant contributions to this field. At this point, artificial neural networks have emerged as a powerful tool in forecasting models, offering flexibility and the capacity to operate without strict prerequisites (Guneri et al., 2012). Additionally, a time series is a sequence formed by arranging observed values of any event in chronological order (Kaynar and Tastan, 2009). Time series analysis, on the other hand, is a method that aims to model the stochastic process that provides the structure of the observed series and make predictions about the future using past observations (Kaynar and Tastan, 2009). These methods allow for the careful analysis of past data and the identification of future trends (Benli and Yıldız, 2015). Among the most frequently used methods in time series analysis are simple exponential smoothing, Holt's linear trend method, and ARIMA models introduced by Box and Jenkins (Benli and Yıldız, 2015). There are numerous studies in the literature demonstrating the success of feedforward ANN models for time series forecasting (Guler et al., 2017). The flexible structure and data processing capabilities of ANNs provide significant contributions to time series analyses (Allende et al., 2002). Deep learning-based approaches such as LSTM, GRU, and transformer are commonly applied in time series forecasting (Alizadeh and Nourani, 2024; Xiao et al., 2024; Si et al., 2024).

2.2.1 DeepDenT Application

Deep networks such as CNN, LSTM and GRU are commonly used because of their modular nature that can be constructed easily using libraries. However, in recent history, shallow ANNs implemented using different models of neurons and architectures are being found to result in more favorable prediction results over deep networks. The deep dendritic artificial neural network (DeepDenT) used in this study develops a new dendritic cell structure that can process multiple inputs and outputs such as LSTM and GRU cells (Egrioglu and Bas, 2024).

The model employed in our study is based on the DeepDenT-based automatic forecasting method proposed by Bas and Egrioglu (2024). In this method, the structure of the lagged variables (univariate or multivariate) is automatically determined by the algorithm with the help of partial autocorrelation coefficients and input significance tests. As stated in Steps 2 and 4, the model considers all statistically significant lagged variables during the forecasting process. Therefore, the model does not rely solely on a single-lag assumption. All meaningful structures, including multiple lags, were automatically evaluated by the algorithm.

In the study, the DeepDenT application automatically performs the forecasting process after splitting the data into training (80%), validation (10%), and test (10%) sets. The model is trained using the training data and validated using the validation set. Subsequently, the trained model is evaluated on the test set to obtain the final test results. The application provides only the time series error metrics for the test data as output. The software processes the test data based on the model that yields the best performance on the training and validation sets. The DeepDenT selects the best-performing model according to its validation performance and conducts forecasting accordingly. The algorithmic steps of the DeepDenT application are as follows (Bas and Egrioglu, 2024):

Algorithm 1. Automatic “DeepDenT” Algorithm

Step 1. The stationarity of the time series is examined. If the time series is found to be non-stationary, stationarity is achieved through differencing. Initially, it is assessed whether the non-stationarity arises from seasonality. If the following condition holds, the series is seasonally differenced using Equation (1).

$$|ACF_m| > 1.645\sqrt{\frac{1+2(ACF_1+\sum_{i=2}^{m-1} ACF_i^2}{n}} \tag{1}$$

$$Z_t = (1 - B^s)^D x_t \tag{2}$$

In Equation (2), s represents the seasonal period. After examining the non-stationarity caused by seasonality, non-stationarity stemming from trend is analyzed using unit root tests. The augmented Dickey-Fuller test is performed on the time series. If a unit root is detected, differencing is carried out using Equation (3).

$$Z_t = (1 - B)^d Z_t \tag{3}$$

Step 2. The lagged variables to be utilized in modeling the time series with DeepDenT are determined using the partial autocorrelation coefficients computed for the series and the corresponding standard errors of these coefficients.

Step 3. The deep neural network is trained using the lagged variables identified in Step 2.

Step 4. Important lagged variables are identified by performing an input significance analysis on the "DeepDenT" artificial neural network.

Step 5. "DeepDenT" is executed using the significant lagged variables identified in Step 4, and the process returns to Step 3.

Step 6. Predictions are generated by the trained **"DeepDenT"** model and then inverse-transformed based on the pre-processing steps applied earlier.

2.3. Artificial Neural Networks

Artificial neural networks can be defined as structures inspired by the human brain, consisting of weighted connections that are linked together and perform parallel information processing (Ozkan, 2011). In other words, ANNs are computer systems developed to autonomously and automatically perform the learning and discovery capabilities of the human brain (Guner et al., 2022). Moreover, ANNs can be considered the most powerful problem-solving technique capable of processing uncertain, incomplete, and non-normally distributed information (Zorlutuna and Bercan, 2019).

Figure 1 shows a typical artificial neuron, which consists of inputs, weights, a summation function, an activation function, and outputs.

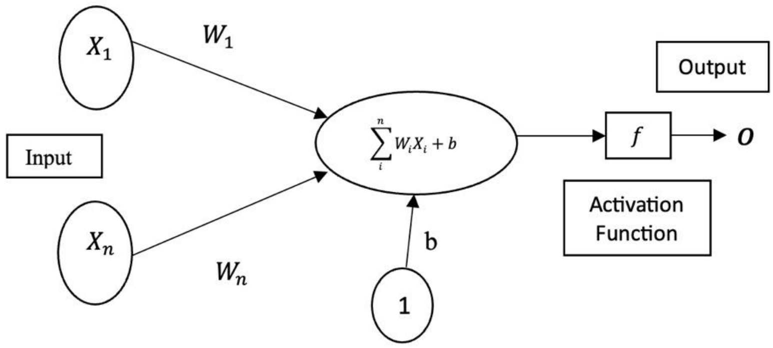

Figure 1. Artificial neural cell.

Time series forecasting with ANNs is used in many fields, and the success of forecasting performance is crucial for implementing the correct policies. This importance has led to increased interest and diversification of forecasting models (Altan, 2008). The six basic steps to be followed when conducting forecasting using ANNs with time series data are as follows (Guneri et al., 2012):

Step 1. Structuring the data: This involves splitting the data into validation, training, and test sets, filling in missing values, and organizing the data in an appropriate format.

Step 2. Preprocessing the dataset: Primarily, the data is scaled to the range [0,1] through transformation as shown in Equation (4). Here, X_i represents the input values.

$$x_i' = \frac{X_i - Min(X_i)}{Max(X_i) - Min(X_i)} \tag{4}$$

Step 3. Modeling: This is the process where the optimal algorithm is selected for data analysis, and this algorithm is trained on the data. In this step, the model learns from the data to create a function, which is then evaluated for accuracy using the test data. The model's accuracy is achieved by minimizing errors, and cross-validation techniques are typically used to test the model's generalization capability.

Step 4. Calculating the optimal weight values: When creating the ANN model, the learning algorithm described in Step 3 is run on the training data to find the optimal weight values. Using these weights, the model's output predictions are then calculated.

Step 5. Evaluating performance metrics: The ANN model's predictions (forecasts) on the test data are produced. Based on the differences between the actual values of the test data and the predicted values, the specified performance metric is calculated.

Step 6. Forecasting: Using the optimal weight values determined in Step 4, future predictions are generated.

3. Findings and Discussion

The study analyzes Türkiye's tobacco production, area harvested, and yield data between 1961 and 2022. Using artificial neural networks, the future levels of production, area harvested, and yield are forecasted. The data were obtained from reliable institutions such as FAO and TurkStat. The study projects a decline in both production areas and area harvested, but an increase in yield. Initially, the performances of time series forecasting models applied to production, area harvested, and yield were listed using various error metrics. In addition, the linear Granger causality test was applied to examine whether there are statistically significant linear relationships among the variables - production, area harvested, and yield - based on the RMSE values provided by the DeepDenT application. The test results, as shown in Table 4, indicated no significant linear causality between these variables at the 5% significance level. This suggests that past values of one variable do not significantly help in forecasting another within this model framework, supporting the independence of variable-specific forecasting structures.

Significant fluctuations occurred in Türkiye's tobacco policies between 1961 and 2022, with supply and demand shocks affecting production, area harvested, and yield during different periods. This study analyzes Türkiye's tobacco production processes using artificial neural networks and provides forecasts for the 2023-2027 period. These forecasts will be valuable for sustainability and planning in tobacco production.

As a result of our time series analysis, the fluctuations in tobacco production over the years and the possible causes of these fluctuations have been examined. The forecasts we made using the time series model provide important insights into future tobacco production and yield levels. In Table 2, the forecast results for the years were calculated and listed using the DeepDenT application, a freely accessible deep dendritic artificial neural network MATLAB application (Egrioglu ve Bas, 2024). According to the forecasts presented in Table 2, a gradual decrease in tobacco production and area harvested is projected between 2023 and 2027, while an increase in yield is expected. This suggests that despite the decline in tobacco production and the area allocated for cultivation, yield may increase due to effective agricultural policies and advancing technology.

Table 2. Forecasted values for tobacco production, harvested area, and yield in Türkiye (2023–2027), based on DeepDenT model.

Forecasting of Years	Production	Area Harvested	Yield
2023	80471.35	68877.08	11111.07
2024	78374.09	62703.19	11203.65
2025	76699.89	56528.57	11253.34
2026	74628.92	50354.66	11281.32
2027	72637.27	44180.76	11325.71

In Figure 2, the annual values of tobacco data for production, area harvested, and yield between 1961 and 2022, as well as the forecast results for 2023 to 2027, are presented in a line graph. The forecast results, shown in red in the graph, predict a decrease in production and area harvested, while yield is expected to increase.

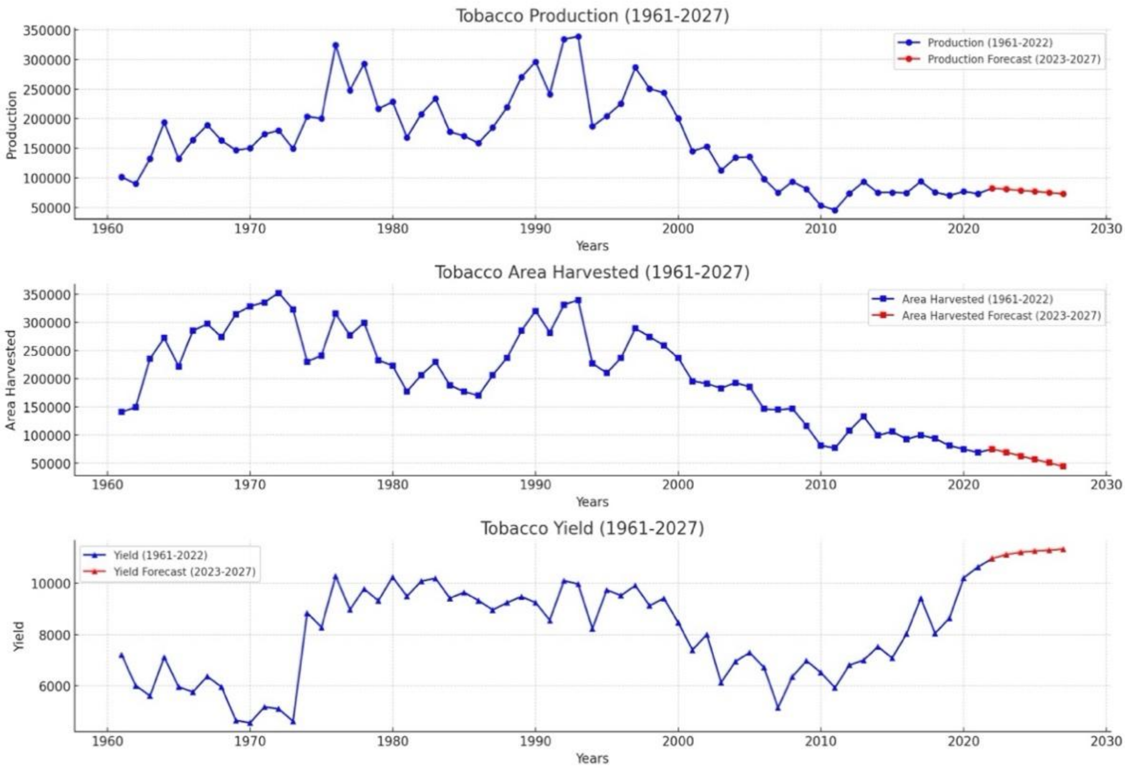

Figure 2. Annual values (1961–2022) and five-year forecasts (2023–2027) for tobacco production, harvested area, and yield in Türkiye.

As a result of these analyses, Table 3 lists the performances of the time series forecasting models applied to tobacco production, area harvested, and yield, using various error metrics such as root mean squared error, mean absolute percentage error (MAPE), symmetric mean absolute percentage error (SMAPE), mean absolute scaled error (MASE), and relative inverse mean absolute error

(REIMAE). It is observed that the area harvested exhibits the lowest values in both absolute MASE and MAPE rates and shows the best performance in error metrics compared to the other two variables. In contrast, the model appears to be more prone to errors in the forecasts for production. Although the yield forecasts are generally good, it is evident that there are still areas for improvement. The forecasting model provides significant improvements over basic models, but more advanced methods may be needed, particularly to minimize errors in production and yield forecasts.

Table 3. Performance metrics of the DeepDenT model for forecasting tobacco production, harvested area, and yield in Türkiye (test results for 2017–2022).

	RMSE	MAPE	SMAPE	MASE	REIMAE
Prdouction	12528.15	0.9581698	0.3886642	0.190768	2.034409
Area Harvested	7870.927	0.735053	0.1973015	0.12322	4.498653
Yield	1064.195	0.8663718	0.3882123	0.660289	0.959894

Table 4 presents the results of the linear Granger causality test. This test investigates linear causal relationships between time series. If the p-values in the Granger statistics are above 0.05, it indicates that one variable does not have a linear effect on the other, suggesting no direct relationship between these variables. According to Table 4, there is not statistically significant (linear) causal relationship between production, area harvested, and yield. These results indicate that the past values of one variable do not have a meaningful impact on predicting the future values of another. This suggests that external factors (e.g., climatic conditions, market dynamics) are more decisive in tobacco production.

Table 4. Results of the linear Granger causality test performed in R for tobacco production, harvested area, and yield.

	F Test	P-Value
Production / AreaHarvest	0.9657	0.3870
Area Harvested / Production	1.1635	0.3199
Yield / AreaHarvest	1.3180	0.2759
Area Harvested / Yield	0.6030	0.5507
Yield / Production	1.0144	0.3692
Production / Yield	1.1722	0.3172

4. Conclusions and Recommendations

The results of this study provide valuable insights into the trends and dynamics of tobacco production in Türkiye. Based on tobacco production, area harvested, and yield data between 1961 and 2022, the forecasts for 2023-2027 in Table 2 were calculated using single-lag time series models. According to the forecast results, a decrease in production and area harvested is expected, while an increase in yield is projected. Table 3 presents the forecasting model results for production, area harvested and yield, validated through several error measures. Metrics such as RMSE, MAPE, SMAPE, MASE, and REIMAE offer complementary insights into the accuracy and reliability of the forecasts. Lower values in these metrics indicate better model performance. Among the three variables, the most accurate predictions were obtained for harvested area, while production forecasts exhibited relatively higher error rates.

In the future, more advanced ANN models, such as complex networks (CN), are planned to be used to make tobacco production forecasts more accurate and reliable (Mata, 2020). CN is an approach that, through complex and multi-layered structures, can model nonlinear relationships more effectively and has been successful with larger datasets. ANN models have effectively captured the nonlinear patterns in time series data, particularly for the area harvested and yield. However, production forecasts exhibited higher errors, suggesting that more complex or refined models may be necessary to improve accuracy in this area.

Limitations of the DeepDenT tool used in this study are some of the main limitations. The software does not report error rates for training and validation sets automatically when training the model, thus making it difficult to quantify the model's capacity to generalize and to detect potential overfitting. Moreover, the software lacks the functionality for manual adjustment of the lag structure. This limitation restricts the researcher’s flexibility in defining the model’s architecture. It is hard to reproduce the research owing to such limitations and reduces methodological transparency.

As shown in Table 4, the results of the linear Granger causality test conducted between production, area harvested, and yield indicate that there is no statistically significant linear causality among these variables at the 5% significance level. According to the test results, no direct linear causality relationship was found among the three variables. This suggests that these variables may be influenced by external factors not included in the current models (e.g., climatic conditions, market dynamics, agricultural policies).

Overall, the findings emphasize the potential of ANN and machine learning models in agricultural forecasting, while also underscoring the need for continuous refinement and exploration of additional factors that may influence production trends. This information can assist policymakers and stakeholders in developing strategies to improve land use efficiency and agricultural yield in the tobacco sector. Future studies should consider incorporating external economic and environmental variables to improve model accuracy and provide a more comprehensive understanding of the factors affecting tobacco production.

Acknowledgements

The authors would like to thank Assoc. Prof. Melih AĞRAZ and Prof. Dr. Erol EĞRİOĞLU for their valuable guidance and contributions throughout this study.

Authors' Contributions

All authors jointly contributed to the conception, data analysis, model development, and manuscript preparation.

Statement of Conflicts of Interest

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The author declares that this study complies with Research and Publication Ethics.

References

Agraz, M., Goksuluk, D., Zhang, P., Choi, B. R., Clements, R., Choudhary, G., and Karniadakis, G. (2024). ML-GAP: Machine Learning-Enhanced Genomic Analysis Pipeline with Autoencoders and Data Augmentation. Frontiers in Genetics, 15, 1442759.

Agraz, M. (2024). Evaluating single multiplicative neuron models in physics-informed neural networks for differential equations. *Scientific Reports*, 14(1), 19073.

Alizadeh, M. J., and Nourani, V. (2024). Multivariate GRU and LSTM models for wave forecasting and hindcasting in the southern Caspian Sea. *Ocean Engineering, 298*, 117193.

Allende, H., Moraga, C., Salas, R. (2002). Artificial neural networks in time series forecasting: A comparative analysis. *Kybernetika*, 38(6), 685-707.

Alpaslan, F., Egrioglu, E., Aladag, C. H., Tiring, E. (2012). Forecasting the lowest and highest gold prices using artificial neural networks. *Journal of Statistical Research*, 9(2), 12-19.

Altan, S. (2008). An alternative approach for forecasting exchange rate performance using artificial neural networks. *Journal of Gazi University Faculty of Economics and Administrative Sciences*, 10(2), 141-160.

Anonymous, (2021). Food and Agriculture Organization of the United Nations. Available from: http://www.fao.org/faostat/en/#data/QC.

Anonymous, (2022). Turkish Statistical Institute. Available from: https://biruni.tuik.gov.tr/medas/?kn=92&locale=tr.

Ataseven, B. (2013). Forecasting modeling with artificial neural networks. *Journal of Proposal*, 10(39), 101-115.

Asi, E., Gozum, S. (2020). Reactions to policies on tobacco products and their implications in Turkey. *Journal of Anatolian Nursing and Health Sciences*, 23(2), 319-330.

Bas, E., and Egrioglu, E. (2024, September 5–6). *A new automatic forecasting method based on deep dendritic artificial neural network*. The 18th International Days of Statistics and Economics, Prague, Czech

Republic, pp. 24–31. https://msed.vse.cz/msed_2024/article/msed-2024-778-paper.pdf.

Benli, Y. K., Yıldız, A. (2015). Forecasting gold prices using time series methods and artificial neural networks. *Journal of Dumlupınar University Social Sciences*, (42).

Bilir, N., Cakir, B., Dagli, E., Erguder, T., Onder, Z. (2010). Tobacco control policies in Turkey. World Health Organization Report. Available from: URL: http://www.euro.who.int/documentE93038.

Celik, S. (2020). Estimation modelling of tobacco production in Turkey: comparativ e analysis of artificial neural networks and multiplicative decomposition methods. *International Journal of Trend in Research and Development*, 7(4), 154-187.

Chen, G., Jiang, S., Chen, D. (2024). Enhanced tobacco yield prediction using spatial information and exogenous variable-driven machine learning models. *Journal of Scientific Research and Reports*, 30(9), 2401.

Egrioglu, E., and Bas, E. (2024). A new deep neural network for forecasting: Deep dendritic artificial neural network. *Artificial Intelligence Review*, 57(7), 171.

Erilli, N. A., Egrioglu, E., Yolcu, U., Aladag, C. H., Uslu, V. R. (2010). Forecasting inflation in Turkey using a hybrid approach of feed-forward and feedback artificial neural networks. *Journal of Dogus University*, 11(1), 42-55.

Guler, D., Saner, G., Naseri, Z. (2017). Forecasting the import volumes of oilseed crops using ARIMA and artificial neural network methods. *Journal of Balkan and Near Eastern Social Sciences*, 3(1), 60-70.

Gumus, A. H., Gumus, S. G. (2005). Current practices and future of the tobacco sector in Turkey. *Journal of Agricultural Economics,* 11(1 and 2), 81-89.

Guneri Tosunoglu, N., Keskin Benli, Y. (2012). Forecasting of Morgan Stanley Capital International Turkey Index using artificial neural networks. *Ege Academic Review*, 12(4).

Guner, S. N., Demir, H. U. (2022). Forecasting iron and steel imports using artificial neural networks and time series methods. *Journal of Sakarya Economics*, 11(3), 389-397.

Karabacak, K. (2017). Tobacco farming and its geographical distribution in Turkey. *Journal of Geographical Sciences*, 15(1), 27-48.

Kaynar, O., Tastan, S. (2009). Comparison of MLP artificial neural networks and ARIMA models in time series analysis. *Journal of Erciyes University Faculty of Economics and Administrative Sciences*, (33).

K. Sameshima and L. A. Baccala, Methods in Brain Connectivity Inference Through Multivariate Time Series Analysis. Boca Raton, FL, USA: CRC Press, 2016.

Liu, X., and Wang, W. (2024). Deep time series forecasting models: A comprehensive survey. *Mathematics*, *12*(10), 1504.

Magazzino, C., Mele, M., and Schneider, N. (2021). A machine learning approach on the relationship among solar and wind energy production, coal consumption, GDP, and CO2 emissions. *Renewable Energy*, 167, 99-115.

Mata, A. S. D. (2020). Complex networks: a mini-review. *Brazilian Journal of Physics*, 50, 658-672.

Ozkan, F. (2011). An alternative approach to exchange rate forecasting with artificial neural networks. Journal of Eskisehir Osmangazi University, *Faculty of Economics and Administrative Sciences*, 2.

Pesen, S. F., Karadogan, S., Akbulut, A. (2021). A general overview of tobacco use and tobacco control policies in the world and Turkey. *Turkey Health Literacy Journal*, 2(3), 191-196.

Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems,* 28.

Si, Y., Nadarajah, S., Zhang, Z., and Xu, C. (2024). Modeling opening price spread of Shanghai Composite Index based on ARIMA-GRU/LSTM hybrid model. *PLOS ONE, 19*(3), e0299164.

Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F., and Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable energy,* 105, 569-582.

Xiao, J., Bi, S., and Deng, T. (2024). Comparative analysis of LSTM, GRU, and Transformer models for stock price prediction. *Manuscript submitted for publication*.

Zhang, M., Wang, J., Li, X. (2020). Tobacco yield prediction and disease detection using hyperspectral imaging and machine learning methods (BPNN, RF). *Journal of Agricultural Engineering and Remote Sensing*, 35(2), 123-135.

Zorlutuna, S., Bircan, H. (2019). Comparison of time series analysis and artificial neural network methods in forecasting the number of tourists coming to Turkey. *Journal of Cumhuriyet University Faculty of Economics and Administrative Sciences*, 20(2), 164-185.