

Article

Modeling the CO₂ Emissions of Turkey Dependent on Various Parameters Employing ARIMAX and Deep Learning Methods

Dilek Surekci Yamacli ^{1,*}  and Cagatay Tuncsiper ²
¹ Department of Economics, Izmir Democracy University, Izmir, 35140, Turkey

² Centrade Fulfillment Services Ltd., Izmir 35010, Turkey; tuncsipercty2@gmail.com

* Correspondence: dilek.surekciyamacli@idu.edu.tr; Tel.: +90-5332524948

Abstract: CO₂ emission is a big problem, not only for current living beings, but also for nature and the upcoming generations. Therefore, modeling CO₂ emissions is the first step in reducing these emissions. In this work, the CO₂ emissions of Turkey are modeled depending on the gross domestic product, amount of hydroelectric electricity generated, amount of energy generated from coal and the electricity obtained from natural gas power plants. The conventional autoregressive integrated moving average with exogenous variables (ARIMAX) and nonlinear deep learning methods are utilized to model CO₂ emissions for the period of 1960–2020 using yearly data obtained from official sources. The modeling results of the ARIMAX and the deep learning methods are quantitatively assessed using four key figures of merit, namely the coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE) and the root mean square error (RMSE). Considering the coefficient of determination and the other performance parameters, it is observed that the deep learning model provides better performance compared with the ARIMAX model in modeling the annual CO₂ emission data, especially in the pandemic period of 2019–2020. The results show that both the conventional ARIMAX and the nonlinear deep learning methods can be utilized to model CO₂ emissions, therefore providing a crucial step for reducing CO₂ emissions and the carbon footprint.

Keywords: CO₂ emissions; nonlinear modeling; ARIMAX; deep learning



Citation: Yamacli, D.S.; Tuncsiper, C. Modeling the CO₂ Emissions of Turkey Dependent on Various Parameters Employing ARIMAX and Deep Learning Methods. *Sustainability* **2024**, *16*, 8753. <https://doi.org/10.3390/su16208753>

Academic Editors: Andrea Nicolini, Roberto Mancinelli and Gang Li

Received: 5 August 2024

Revised: 29 August 2024

Accepted: 1 October 2024

Published: 10 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Environmental quality is an important indicator of the development and civilization status of countries. The forerunner of environmental pollution is greenhouse gas emissions, where CO₂ emissions have the largest share. Greenhouse gases are emitted into the atmosphere as a result of economic production processes, and then these gases cause the greenhouse effect, which leads to climate changes and the increment in average temperatures. The relationship between economic income and environmental pollution is generally modeled using the hypothesis called the environmental Kuznets curve (EKC). This hypothesis expresses that the level of environmental degradation in underdeveloped and developed countries is low, while the pollution caused by developing countries is higher. In fact, this hypothesis is derived from the original Kuznets curve hypothesis, which explains that economic inequality has an inverse U-shaped curve dependent on income per capita [1]. The environmental Kuznets curve hypothesis was proposed by Grossman and Krueger in 1991 [2]. From this viewpoint, it is vital to assess the environmental Kuznets curve analysis of Turkey, as it is a rapidly developing country with a high level of economic growth [3,4]. The original environmental Kuznets curve hypothesis includes economic growth as the independent variable and the level of environmental degradation as the dependent variable, while an inverted U-shaped curve is expected between these variables [2].

Our study delves into the critical issue of environmental degradation, specifically focusing on CO₂ emissions as a primary indicator of this phenomenon from a sustainability

viewpoint. The background of this study is grounded in the growing global concern over climate change and the detrimental effects of greenhouse gas emissions, which are predominantly driven by economic activities. Over the years, numerous studies have highlighted the strong correlation between economic growth, energy production and environmental degradation, particularly CO₂ emissions. However, the complexity of these relationships, especially the nonlinear interactions between economic variables and environmental outcomes, necessitates more sophisticated modeling approaches, as given in detail in the following literature survey section. The rationale for this study stems from the need to improve predictive models that can more accurately capture these relationships, thereby providing valuable insights for policymakers. By comparing the performance of conventional statistical models like ARIMAX with advanced deep learning methods, the authors aim to determine the most effective approach for modeling CO₂ emissions. This is crucial for devising strategies to mitigate the impact of economic activities on the environment and to contribute to the global effort to reduce greenhouse gas emissions.

The aim of this study is to model the level of environmental degradation using economic variables. CO₂ emissions are taken as the indicator for the level of environmental degradation, while the independent variables are selected as the gross domestic product per capita and the electricity generated using coal, hydroelectric power plants and natural gas. The conventional autoregressive integrated moving average with exogenous variables (ARIMAX) and nonlinear deep learning methods are utilized to model CO₂ emissions dependent on the gross domestic product per capita and the electricity generated using coal, hydroelectric power plants and natural gas. Furthermore, the figures of merits of these models such as the coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE) and the root mean square error (RMSE) are computed. The modeling results show that CO₂ emissions can be accurately modeled by both the ARIMAX and deep learning methods, which is significant in understanding the factors affecting CO₂ emissions to reduce greenhouse gases.

2. Literature Survey

The uncontrolled greenhouse gas emissions from industrial production and oil-based vehicles increase the level of environmental degradation. This process was accelerated by the rapid industrialization of countries after World War II, which led to discussions of the scientific evidence for climate changes at the United Nations General Assembly in 1968 and then the Stockholm Conference in 1972, which aimed to address environmental problems with international collaboration [5]. Then, the climate change effects and the potential prevention measures were discussed in detail in 1987, and the report named *Our Common Future* was published by the Brundtland Commission [6]. Furthermore, the Kyoto Protocol was signed by 84 countries in 1997; however, the results of this protocol were limited because only 46 countries ratified the protocol, including the European Union countries and Japan, while industrialized countries such as China, the United States and Australia did not ratify the protocol [7]. Finally, the Paris Agreement, which is a legally binding international treaty to combat climate change, was signed by 196 countries at the UN Climate Change Conference in 2015. The Paris Agreement came into force on 4 November 2016 and aims to hold “the increase in the global average temperature to well below 2 °C above pre-industrial levels” and pursue efforts “to limit the temperature increase to 1.5 °C above pre-industrial levels”. [8]. Turkey has also ratified the Paris Agreement and aims to reduce its total greenhouse gas emissions from the projected 1175 Mt (million tons) to 929 Mt by 2030 [9]. This corresponds to a 21% reduction in total greenhouse gas emissions. It is obviously important to model the greenhouse gas emissions to achieve accurate predictions for the future. Therefore, it is imperative to analyze the main causes and their share in the greenhouse emissions. The studies existing in the literature indicate that the biggest contributor to CO₂ emissions of Turkey is the economic expansion, in other words the scale effect, caused by the increase in the energy production [10]. This is in line with other territories such as the European Union for which it is reported that the industrial processes

and the energy production are named as the source of the greenhouse gas emissions [11]. Hence, it is possible to model the greenhouse gas emissions, particularly CO₂ emissions since it has the largest share dependent on the economic growth and the consumption of the fuels for energy production. This modeling not only helps to understand the contributions of various factors to CO₂ gas emissions but also enables us to forecast CO₂ emissions, which aids the policies for the reduction in greenhouse gas emissions according to the Paris Agreement.

There are vast numbers of studies in the literature regarding the assessment of the environmental Kuznets curve hypothesis. In one of these studies, the income per capita and the greenhouse gas emissions for the United States are considered in the period of 1988–1994, where it is shown that the level of environmental degradation decreases as the income per capita increases, which is in line with the environmental Kuznets curve hypothesis [12]. CO₂ emissions and the economic growth of China for the 1991–2006 period are investigated in another study in which it is concluded that the economic scaling has important effects on the level of environmental degradation in China, which is the second country for the greenhouse gas emissions viewpoint [13]. The economies and the air pollution data of 18 selected OECD countries are analyzed in another work employing the panel data analysis method where it is shown that the economic growth and the greenhouse gas emissions data of these countries support the environmental Kuznets curve hypothesis [14]. The empirical study of Bruvoll and Medin also verifies that there is a correlation between the income per capita and the environmental pollution for the developed countries, also verifying the environmental Kuznets curve hypothesis [15]. In another study, it is argued that the environmental degradation increases monotonically in low-income countries and decreases monotonically in high-income countries, while having inverted U characteristics for middle-income countries, as estimated by the environmental Kuznets curve (EKC) hypothesis [16]. The environmental Kuznets curve hypothesis was also studied by Burnett et al. where they have considered the monthly CO₂ emission and income per capita data of the United States for the period of 1981–2003 and they have validated the EKC hypothesis [17]. The causality relationship between CO₂ emission and the economic growth is also studied in the literature in conjunction with the environmental Kuznets curve hypothesis. For example, the economic and the environmental data of Malaysia have been analyzed for the period of 1971–1999 employing the cointegration analysis and it is shown that there is a unidirectional causality relationship for the economic growth and CO₂ emission [18]. Similarly, the data regarding China for the 1981–2006 period are considered employing the cointegration and the Granger causality analysis in which the unidirectional causality relationship from the economic growth to CO₂ emissions was also validated [19]. The cointegration and the Toda–Yamamoto causality analyses are applied on the economic and environmental data of Italy for the period of 1970–2006 and it is indicated that there is a bidirectional causality relationship between the economic growth and the level of environmental degradation [20]. A similar bidirectional causality relationship between the economic growth and CO₂ emissions is shown in a study where the economic and CO₂ data of Tunisia are investigated using Granger causality and a vector error correction model for the 1990–2015 period [21].

The original environmental Kuznets curve hypothesis includes the income per capita as the independent variable; however, some studies in the literature use other related data in addition to the income per capita to increase the accuracy of the modeling of CO₂ emission. For example, the relationship between the economic growth, greenhouse gas emissions and the energy consumption of China, India, Iran, Indonesia and South Africa is investigated using panel data analysis in a recent study [22]. Similarly, in another study, the energy consumption and direct foreign investment data are used in addition to the income per capita for the modeling of CO₂ emission for Vietnam in the period of 1980–2010 [23]. The economies of 65 developing countries are analyzed in another work where the foreign and domestic investments and the gross domestic product are used as the independent variables for the modeling of CO₂ emission employing panel data analysis from the EKC

viewpoint [24]. In another work, CO₂ emissions caused by agricultural processes are modelled dependent on the income per capita and many other data, including direct foreign investments and the share of agriculture in the gross domestic product for selected less-developed countries using panel data analysis [25]. In a similar study, the effects of the income per capita and foreign direct investments on CO₂ emission for China for the period of 1982–2006 are investigated using vector autoregressive analysis [26]. In another study, the environmental Kuznets curve hypothesis is validated for Pakistan for the 1980–2013 period using economic growth, energy consumption, trade openness and population data [27]. In an extensive work including 93 countries, the ecological footprint, economic growth, energy consumption, urbanization and trade openness data are used to assess the environmental Kuznets curve hypothesis for the period of 1980–2008 in which the EKC hypothesis is validated for middle-income and high-income countries [28]. The generalized method of moments, in conjunction with the panel analysis, is employed for the analysis of the air pollution and the economic growth for 14 Asian countries in which the environmental Kuznets curve hypothesis is validated [29]. On the other hand, the autoregressive distributed lag (ARDL) method is utilized for the assessment of the environmental Kuznets curve hypothesis for Malaysia in the period of 1970–2008 where it is shown that the EKC hypothesis is valid [30]. In another study, the ARDL method is also used for the Chinese economy and the environmental data for the 1975–2015 period and this study also validates the environmental Kuznets curve hypothesis [31]. The ARDL approach is also used for the environmental data of Portugal for the 1971–2008 period and the environmental data of Tunisia for the 1971–2010 period where the EKC hypothesis is validated for both countries [32,33]. On the other hand, Granger causality analysis was employed for the analysis of the economic and the environmental data of the United States in the 1960–2007 period considering the relationship among the nuclear energy consumption, renewable energy consumption and the economic growth and it is shown that nuclear energy consumption can be used to reduce the CO₂ emission [34]. In another extensive study, the environmental Kuznets curve hypothesis is assessed for 43 developing countries employing the panel cointegration method for the 1980–2004 period where it is shown that the EKC hypothesis holds for these countries [35]. Similarly, the environmental Kuznets curve hypothesis is used to analyze the economic and environmental data of 13 Eurasian countries using panel data analysis and the panel cointegration method where it is concluded that there exists a bidirectional causality relationship between the energy consumption and the economic growth for these countries [36]. A similar result was achieved for the G7 countries in the period of 1980–2009 using panel ARDL analysis where it is shown that the energy consumption impacts the economic growth [37]. In another work, the economic and environmental degradation data were considered for Portugal for the 1970–2010 period where the relationships among the economic growth, renewable energy consumption, globalization and CO₂ emission are investigated employing several methods such as the generalized method of moments, least squares, vector error correction and the Granger causality analysis and it is shown that CO₂ emission has a causality relationship with the energy consumption, renewable energy and globalization [38]. In another extensive work, the relationship between the renewable energy consumption and the economic growth for 80 selected countries in the period of 1990–2012 was evaluated using a panel error correction model and it is shown that there is a strong causality relationship between the renewable energy consumption and the economic growth [39]. The validity of the environmental Kuznets curve extended with the renewable and non-renewable energy consumption data is tested for 16 European Union countries in the 1990–2008 period using panel data analysis where it is argued that the EKC hypothesis is supported with the result that the utilization of the renewable energy sources helps to reduce the greenhouse gas emissions by about 50% compared to the usage of the non-renewable energy sources [40]. The economic and environmental degradation data of France for the 1960–2000 period are considered in another study employing cointegration and vector error correction methods where it is shown that the increments in the energy use and the economic growth increase

CO₂ emissions in the long run [41]. The structural break cointegration test was utilized for the assessment of the environmental Kuznets curve hypothesis for the Indian economic and environmental data for the 1971–2008 period in which it is shown that the EKC hypothesis is valid for the considered period [42].

The interest in the studies regarding the modeling of CO₂ emission of Turkey has increased due to the ratification of the Paris Agreement by the Turkish government. For example, the utilization of the grey prediction model for the estimation of CO₂ emissions of Turkey dependent on the energy consumption using the data of the 1965–2012 period is investigated in a recent study [43]. Similarly, the causality relationships among the environmental degradation, income per capita and energy consumption are analyzed for Turkey in the 1960–2005 period in another study where it is concluded that the amount of greenhouse gas emissions is affected by the income per capita, energy consumption and the foreign trade in descending order [44]. The relationships among the foreign investments, environmental pollution and the economic growth have been assessed for Turkey for the 1987–2009 period in another study employing the cointegration, error correction and the Granger causality methods where it is shown that foreign investments and the economic growth increases the CO₂ emissions [45]. The relationship among the CO₂ emission, economic growth, electricity consumption and foreign direct investments of Turkey in the 1980–2013 period was evaluated employing cointegration analysis and it is concluded that there is a long-term cointegration relationship between CO₂ emission and the economic growth, electricity consumption and the foreign direct investments [46]. Similarly, the validity of the EKC hypothesis is tested using the economic growth, CO₂ emission and the foreign direct investment data of Turkey for the period of 1974–2011 employing a vector autoregressive approach where it is shown that there exist strong causality relationships among the considered variables [47]. In another study, the autoregressive distributed lag method is employed to investigate the impacts of the economic growth, energy consumption and foreign direct investments on CO₂ emission for Turkey in the 1970–2014 period and it is shown that the economic growth and the energy consumption affect CO₂ emission [48]. The structural break cointegration analysis was employed in another work in which the economic and environmental data of Turkey for the period of 1950–2007 are considered where it is shown that the environmental Kuznets curve hypothesis is valid for Turkey [49]. Similarly, in another study, the structural break cointegration test was utilized to analyze the relationship of CO₂ emission, gross domestic product and the energy consumption of Turkey for the period of 1960–2007 and it is shown that EKC hypothesis is valid for the considered period [50]. The autoregressive distributed lag method with bounds test is used to check the validity of the environmental Kuznets curve hypothesis for CO₂ emission, energy consumption, economic growth, openness and the financial development data of Turkey for the 1960–2007 period in another study where it is shown that the EKC hypothesis is valid [51]. Similarly, the autoregressive distributed lag approach with bounds test was utilized to assess CO₂ emission, economic growth, globalization and the energy consumption data of Turkey for the 1970–2010 period and it is concluded that the environmental Kuznets curve hypothesis is valid for the considered period [52]. The renewable energy consumption data are also considered in some studies using the framework of the EKC hypothesis; therefore, the existence of the causality relationships between the renewable energy consumption and the related data is investigated. For example, the causality relationship between the economic growth and the renewable energy consumption is investigated for Turkey in the 1990–2010 period using the autoregressive distributed lag approach where it is shown that the renewable energy consumption has a negative effect on the economic growth; the Toda–Yamamoto causality test is also used in the same study, which indicates that there exists a unidirectional causality relationship from the economic growth to the renewable energy consumption [53]. In another study, the relationships among the primary energy consumption, economic growth and CO₂ emission data are investigated employing the cointegration and the causality analyses for Turkey in the 1970–2008 period where it is demonstrated that there exist unidirectional

causality relationships from the economic growth to CO₂ emission and also from the energy consumption to CO₂ emission [54].

In this study, the CO₂ emission of Turkey for the period of 1960–2020 is modelled using the autoregressive integrated moving average method with exogenous variable (ARIMAX) and deep learning methods. Independent variables are selected as the income per capita and energy consumptions from the hydroelectric, coal and natural gas sources. The income per capita variable exists in the basic environmental Kuznets curve hypothesis while the energy consumption data from the three main sources are also considered as exogenous variables to increase the accuracy of the modeling of CO₂ emission. The data and the details of the methods employed in the modeling process are explained in the Materials and Method section. The results of the models, the figure of merits of these models and the contribution of these models to the literature are discussed in the Results and Discussion section. The overall results of the work are presented in the Conclusions section.

3. Materials and Methods

The data related to the income per capita, CO₂ emission and energy consumption from hydroelectric, coal and natural sources are taken from the Turkey page of the World Bank and the Our World in Data based on the Global Carbon Project website [55]. As the first step, the gathered data are seasonally adjusted in EViews software using the seasonal-trend decomposition with Loess (STL) method [56]. Then, the augmented Dickey–Fuller (ADF) [57], Phillips–Perron [58] and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) [59] unit root tests are performed for the level and first difference data to analyze the stationarity of the data. Structural break unit root tests are then applied for ensuring stationarity of the data. Results of these tests are given in Table 1. After analyzing and conditioning the time series, models employing ARIMAX and deep learning methods are built.

Two different methods are utilized in this study for modeling the CO₂ emission dependent on the income per capita and energy consumption from three main sources. The first method used for modeling is the autoregressive integrated moving average with exogenous variable (ARIMAX), which enables us to model the dependent variable using its autoregressive terms together with exogenous or explanatory variables. ARIMAX models extend the conventional autoregressive integrated moving average (ARIMA) model with exogenous variables. The ARIMAX method is selected for its effectiveness in modeling CO₂ emissions by incorporating both past values and exogenous variables, providing a balance of robustness and interpretability suited to the objectives of our study. The defining equation of the ARIMAX model for the dependent variable y and exogenous variable x is given in Equation (1) [60].

$$\left(1 - \sum_{s=1}^p \alpha_s L^s\right) \Delta y_t = \sum_{s=1}^q \beta_s L^s x_t + \left(1 + \sum_{s=1}^r \gamma_s L^s\right) e_t \quad (1)$$

In Equation (1), L denotes the lag operator and e_t is the error term. The model parameters, namely μ , α_s , β_s and γ_s , have to be calculated during the modeling process [60–62]. It is worth noting that the ARIMAX model consists of four terms, namely the autoregressive (AR), integrated (I), moving average (MA) and exogenous variable (X) [63]. As in conventional ARIMA models, the ARIMAX model is represented by three integers, p , d and q , such that ARIMAX(p , d , q), where p , d and q are the degrees of the autoregressive, moving average and the integration terms. Eviews software is used for developing the ARIMAX model in this study.

The second method used in this study for the modeling of the CO₂ emission is deep learning networks. Deep learning is a machine learning model where multiple hidden layers are employed to increase the training accuracy [64]. The deep learning approach is selected for its ability to capture complex nonlinear relationships in large datasets, offering a robust and flexible alternative to conventional models, while fully utilizing the available labelled data. In this study, multilayer perceptron (MLP)-type neural networks [65] with

multiple hidden layers are utilized as the deep learning network. Activation functions of the network are selected as tangent hyperbolic functions. The developed deep learning network is shown in Figure 1.

Table 1. Results of ADF, PP and KPSS stationarity tests.

Tests/Variables		CO2SM	GDPPEM	COALSM	HYDROSM	NATURASM	
ADF	Level	Intercept	1.366	2.408	−3.735	−1.519	−1.164
		Prob.	0.998	1.000	0.005	0.517	0.684
		t stat.	−3.544	−3.544	−3.544	−3.544	−3.546
		Trend &Intercept	−1.770	−0.233	−3.617	−2.434	0.162
		Prob.	0.707	0.990	0.036	0.359	0.997
		t stat.	−4.118	−4.118	−4.118	−4.118	−4.118
	1st Difference	Intercept	−6.305 *	−6.579 *		−7.306 *	−4.685 *
		Prob.	0.000	0.000	−	0.000	0.000
		t stat.	−3.546	−3.546		−3.546	−3.546
		Trend &Intercept	−6.578 *	−7.244 *		−7.262 *	−4.759 *
		Prob.	0.000	0.000	−	0.000	0.001
		t stat.	−4.121	−4.121		−4.121	−4.121
PP	Level	Intercept	2.089	3.324	−3.593 *	−1.519	−1.171
		Prob.	0.999	1.000	0.008	0.517	0.681
		t stat.	−3.544	−3.544	−3.544	−3.544	−3.544
		Trend &Intercept	−1.657	−0.197	−2.091 *	−2.434	−0.773
		Prob.	0.757	0.991	0.008	0.359	0.962
		t stat.	−4.118	−4.118	−3.544	−4.118	−4.118
	1st Difference	Intercept	−6.166 *	−6.573 *		−7.298 *	−4.756 *
		Prob.	0.000	0.000	−	0.000	0.000
		t stat.	−3.546	−3.546		−3.546	−3.546
		Trend &Intercept	−6.328 *	−7.232 *		−7.250 *	−4.823 *
		Prob.	0.000	0.000	−	0.000	0.001
		t stat.	−4.121	−4.121		−4.121	−4.121
KPSS	Level	Intercept	0.944	0.918	0.366 *	0.574 ***	0.795
		t stat.	0.739	0.739	0.739	0.739	0.739
		Trend &Intercept	0.245	0.223	0.121 *	0.187 ***	0.123 ***
		t stat.	0.216	0.216	0.127	0.216	0.146
	1st Difference	Intercept	0.419 *	0.584 *		0.079 *	0.241 *
		t stat.	0.739	0.739	−	0.739	0.739
		Trend &Intercept	0.097 *	0.087 *		0.052 *	0.204 **
		t stat.	0.216	0.216	−	0.216	0.246

*: 1%, ** 5%, *** 10% levels.

As the first step, the deep learning network shown in Figure 1 is trained for optimizing the coefficients of the neurons of the network. A total of 70% of the available data are used for the training phase, while the remaining 30% of the data are employed as the test data to test the accuracy of the network. Nonlinear activation functions of the neurons are selected as hyperbolic tangent functions in the deep learning network. The presented deep learning network is developed in the Python programming language using the NumPy and Pandas numerical libraries and the MLPRegressor class from the SciKit-Learn (SKLearn) library [66–68]. The modelling results are given in the following Results and Discussion section.

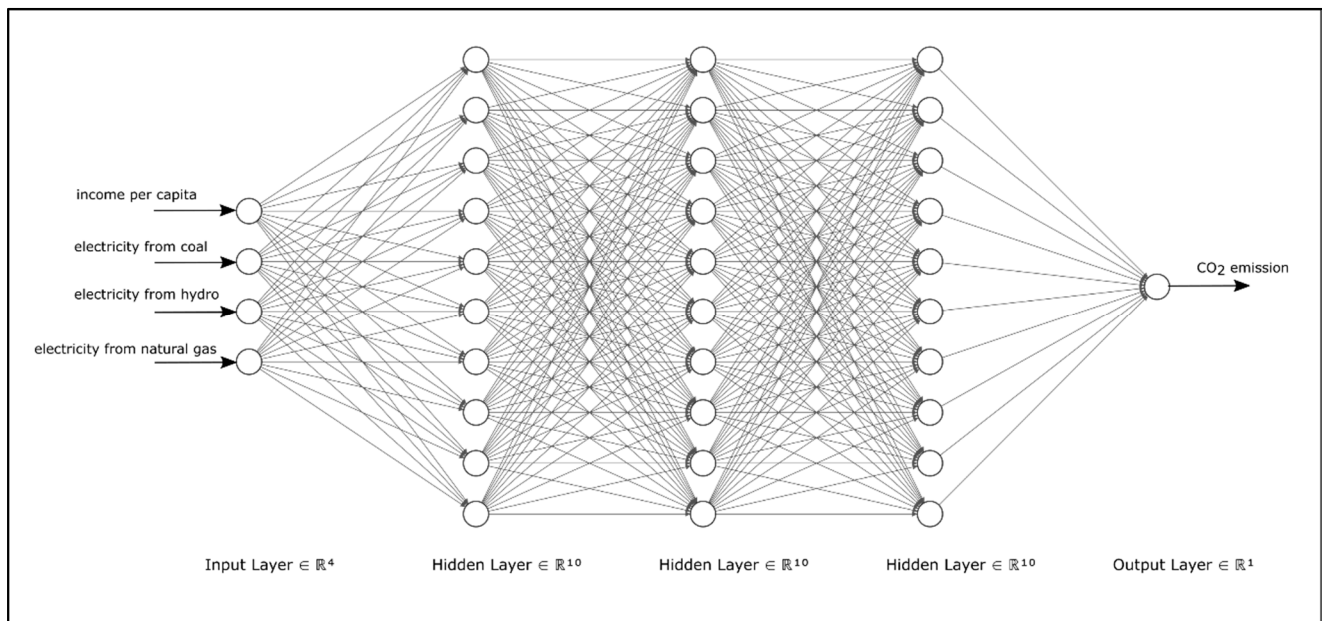


Figure 1. Structure of the developed deep learning network.

4. Results and Discussion

The annual CO₂ emission data of Turkey in the period of 1960–2020 are taken from the Our World in Data based on the Global Carbon Project website [69]. These CO₂ data per year are plotted as in Figure 2 using the Matplotlib library in Python [70].

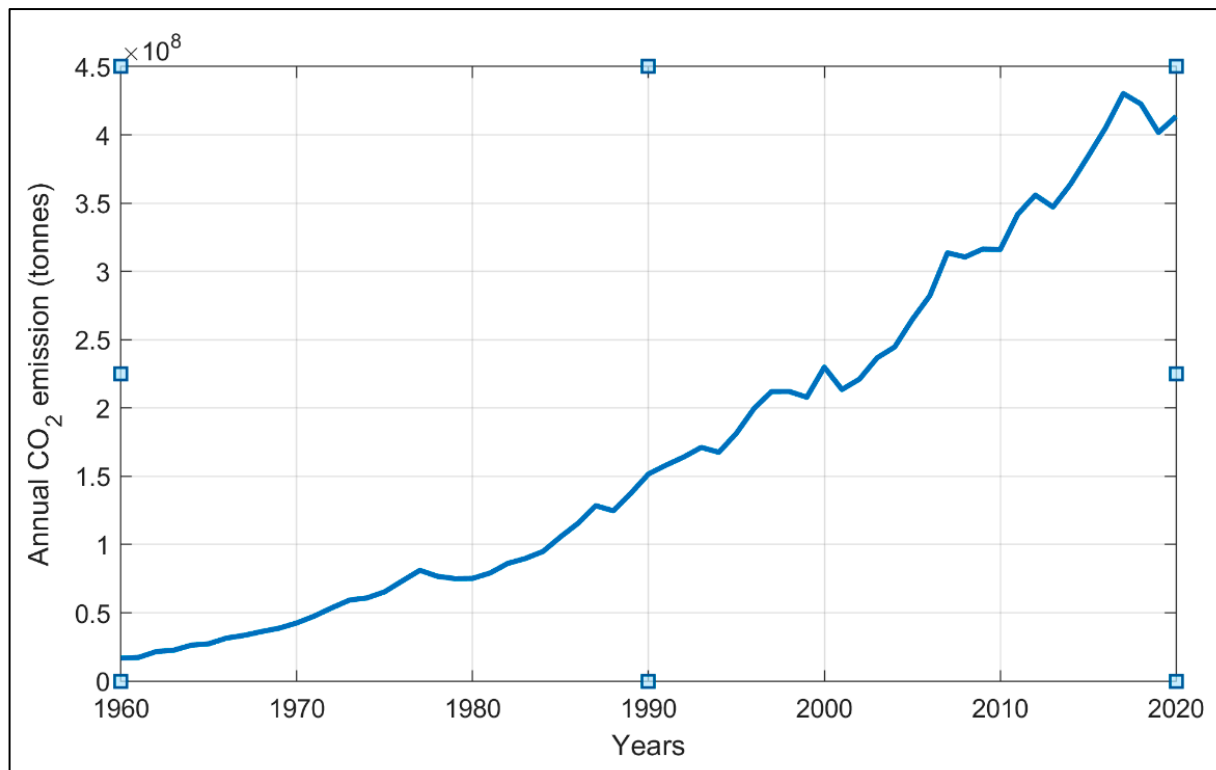


Figure 2. The annual CO₂ emission data of Turkey for the period of 1960–2020.

Firstly, Eviews software is utilized for the automatic ARIMAX modeling of the CO₂ data dependent on the income per capita and the electricity consumption from hydroelectric, coal and natural gas resources. Structural break unit root tests are performed before the

actual ARIMAX modeling phase. The results of the unit root tests are given in Table 1. The results of the automatic ARIMAX modeling are also summarized in Table 2.

Table 2. ARIMAX modeling summary.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COALSM	306.3298	24,953.65	0.012276	0.9903
D(GDPPEM)	3702.906	3406.551	1.086996	0.2839
D(HYDROSM)	351,132.6	215,727.8	1.627665	0.1119
D(NATURASM)	127,468.8	105,479.1	1.208474	0.2343
@MONTH = 1	−310,626.9	953,166.0	−0.325890	0.7463
AR(1)	2.000207	0.236319	8.464029	0.0000
AR(2)	−1.966760	0.571676	−3.440342	0.0014
AR(3)	1.308354	0.628479	2.081779	0.0441
AR(4)	−0.586621	0.293817	−1.996557	0.0531
SAR(1)	−1.631272	0.214531	−7.603889	0.0000
SAR(2)	−0.770918	0.145517	−5.297800	0.0000
MA(1)	−1.999882	9.345006	−0.214005	0.8317
MA(2)	0.999882	9.333498	0.107128	0.9153
SMA(1)	−0.138304	1279.690	−0.000108	0.9999
SMA(2)	−0.861696	10,348.26	−8.33E−05	0.9999
SIGMASQ	2.91E+13	1.97E+16	0.001478	0.9988
R-squared	0.823937	Mean dependent var		234,617.2
Adjusted R-squared	0.754438	S.D. dependent var		12,975,655
S.E. of regression	6,429,984.	Akaike info criterion		34.92904
Sum squared resid	1.57E+15	Schwarz criterion		35.51837
Log likelihood	−927.0840	Hannan–Quinn criter.		35.15632
Durbin-Watson stat	1.994260			

As the next step, the deep learning modelling is performed. A total of 70% of the available data are used as the training data and the remaining 30% of the data are used as the test data, which is a standard data split procedure. The `train_test_split` class of the SKLearn library is utilized for splitting the available data. The loss curve associated with the training phase of the developed loss curve is given in Figure 3.

As can be seen from Figure 3, the training phase of the deep learning network is successfully completed, minimizing the loss value. As the next step, actual CO₂ emission data, CO₂ emission data obtained from the ARIMAX model and the CO₂ emission data as the result of the developed deep learning network are plotted on the same axes in Figure 4. As can be observed from Figure 4, both the ARIMAX model and the deep learning network accurately model the annual CO₂ emission data of Turkey in the 1960–2020 period employing the income per capita and the electricity consumption from the hydroelectric, coal and natural gas resources. In order to objectively assess the performance of the ARIMAX and deep learning models, figure of merit values, namely the coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE) and the root mean square error (RMSE), are computed in Python using the formulas given in Equations (2)–(5) [71].

$$R^2 = \frac{\sum_1^{dim} (O - avg(O))^2 - \sum_1^{dim} (O - N)^2}{\sum_1^{dim} (O - avg(O))^2} \quad (2)$$

$$MAE = \frac{\sum_1^{dim} |O - N|}{dim} \quad (3)$$

$$MAPE = \frac{100}{dim} \sum_1^{dim} \left| \frac{O - N}{N} \right| \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_1^{dim} (O - N)^2}{dim}} \quad (5)$$

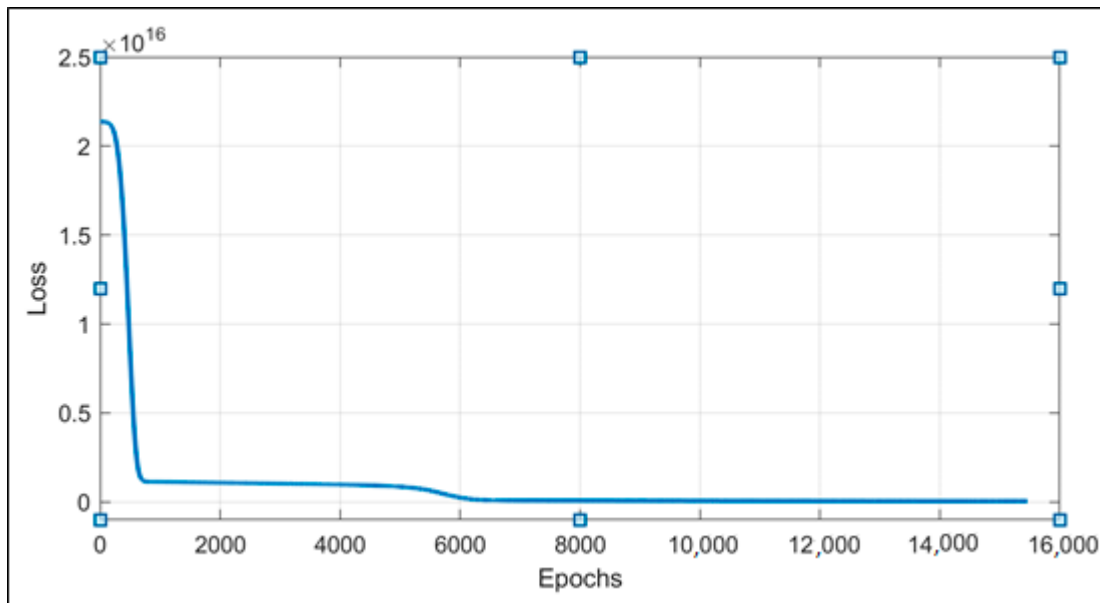


Figure 3. The loss curve of the training phase of the deep learning network.

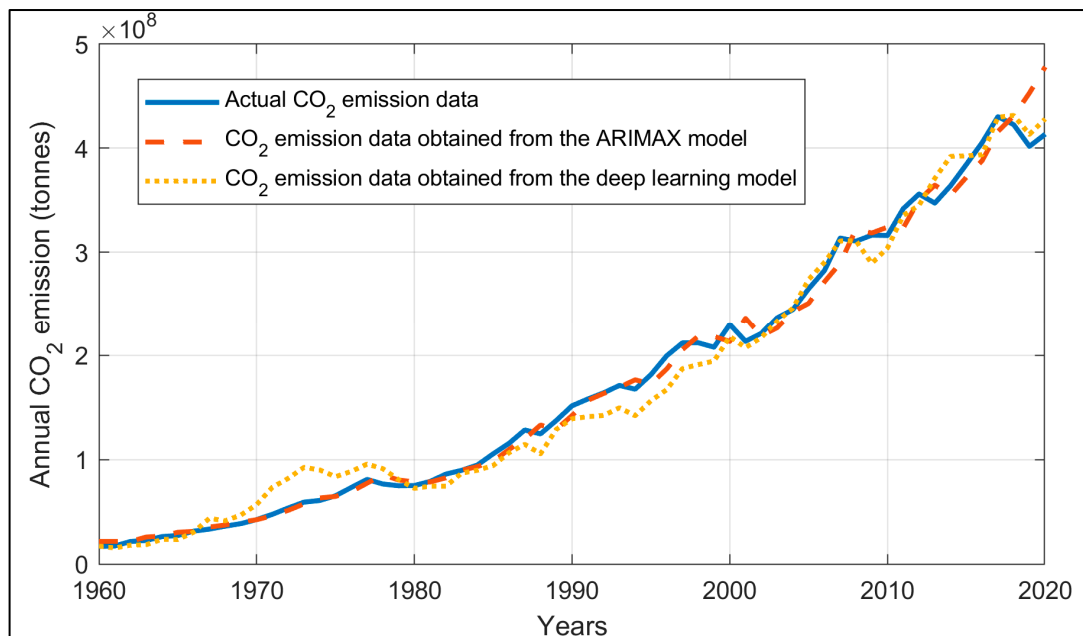


Figure 4. The actual annual CO₂ emission data and the CO₂ emission data obtained from the ARIMAX and deep learning models.

In Equations (2)–(5), O represents the original data, N is the data obtained from the model and dim is the length of the time series. R², MAE, MAPE and RMSE values for the

ARIMAX and deep learning model results are calculated in the Python environment using the classes of the SKLearn.Metrics library, as presented in Table 3.

Table 3. Performance parameters of the ARIMAX and the deep learning models.

Model	R ²	MAE	MAPE	RMSE
ARIMAX model	0.82	9,748,530	0.079	28,788,426
Deep learning model	0.98	8,574,162	0.055	11,133,153

As can be seen from Table 3, both ARIMAX and deep learning methods can be utilized to model the annual CO₂ emission data dependent on the income per capita and electricity consumption from coal, hydroelectric and natural gas resources for Turkey in the period of 1960–2020. Considering the coefficient of determination and other performance parameters, it is observed that the deep learning model provides better performance compared to the ARIMAX model for modelling the annual CO₂ emission data especially in the pandemic period of 2019–2020. It can be argued that either the ARIMAX model or the deep learning model can be used to model the annual CO₂ emission data, providing insights for the estimation of the future emission data dependent on the income per capita and electricity usage, which would help to achieve the goals regarding the reduction in the carbon footprint according to the Paris Agreement.

5. Conclusions

Accurate modeling of annual CO₂ emission data is of crucial importance considering the recently ratified Paris Agreement. Therefore, annual CO₂ emission of Turkey for the period of 1960–2020 is modelled in this work. According to the original environmental Kuznets curve hypothesis, income per capita is the independent variable while CO₂ emission is the dependent variable. In order to increase the performance of modelling the annual CO₂ emission data, additional independent variables, namely the electricity consumption obtained from hydroelectric, coal and natural gas resources, are also included in this study. Two different approaches, ARIMAX and deep learning methods, are utilized for modelling the annual CO₂ emission data dependent on income per capita and electricity consumption from the hydroelectric, coal and natural gas resources. Firstly, annual CO₂ emission data and the independent variables are seasonally adjusted in EViews software using the STL method. Structural break unit root tests are then applied for ensuring the stationarity of the data. Results of these tests are given in Table 1. Then, the ARIMAX model is developed in the EViews software. In the second part of the modeling phase, a deep learning network, consisting of three hidden layers employing multilayer perceptron neural networks, is developed in the Python programming language using the SciKit-Learn library. The available data are split as training and test data using the test_train_split class of Python, where 70% of the data are taken as the training data and the remaining 30% of the data are taken as the test data according to the standard procedure.

Actual annual CO₂ emission data and CO₂ emission obtained from the ARIMAX model and deep learning model are plotted on the same axes to visually inspect the performance of the developed ARIMAX and deep learning models. The plotted data indicate accurate modeling performance. In addition, the performance parameters of the ARIMAX and deep learning models, namely coefficient of determination (R²), mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE) values, are computed in the Python environment using the SciKit-Learn-Metrics library. The performance parameters show that both ARIMAX and deep learning models provide high-performance modelling of the annual CO₂ emission data. Considering the coefficient of determination and other performance parameters, it is observed that the developed deep learning model provides higher accuracy compared to the ARIMAX model for modeling the annual CO₂ emission data especially in the pandemic period of 2019–2020. Obtained results show that the annual CO₂ emission data of Turkey can be modelled using

both conventional ARIMAX and nonlinear deep learning methods, which would aid in forecasting CO₂ emission data in order to make the required policy changes to achieve the aims of the Paris Agreement. In summary, the contributions of this study can be stated as follows. (i) The annual CO₂ emission data are modelled utilizing explanatory variables, namely the electricity consumption obtained from the hydroelectric, coal and natural gas resources, to increase the accuracy, (ii) the conventional ARIMAX method is used to model the CO₂ emission, (iii) a deep learning MLP network is also developed to model the CO₂ emission, (iv) the results of the ARIMAX and deep learning MLP network are compared and (v) it is observed that both ARIMAX and the deep learning MLP network can be used to model the CO₂ emission data; however, the deep learning model provides higher accuracy compared to the ARIMAX model.

Author Contributions: C.T.: Methodology, software, validation, writing—review and editing. D.S.Y.: Conceptualization, data curation, methodology, visualization, writing—original draft, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Dataset available on request from the authors.

Conflicts of Interest: Author Cagatay Tuncsipir was employed by the company Centrade Fulfillment Services Ltd. The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Kuznets, S. Economic growth and income inequality. *Am. Econ. Rev.* **1955**, *45*, 1–28.
2. Grossman, G.M.; Krueger, A.B. *Environmental Impacts of a North American Free Trade Agreement*; Working Paper 3914 of the National Bureau of Economic Research; National Bureau of Economic Research: Cambridge, MA, USA, 1991; Volume 3, pp. 27–32.
3. Benli, M. External debt burden—Economic growth nexus in Turkey. *Soc. Sci. Res. J.* **2020**, *9*, 107–116.
4. Etokakpan, M.U.; Osundina, O.A.; Bekun, F.V.; Sarkodie, S.A. Rethinking electricity consumption and economic growth nexus in Turkey: Environmental pros and cons. *Environ. Sci. Pollut. Res.* **2020**, *27*, 39222–39240. [\[CrossRef\]](#)
5. Report of the United Nations Conference on the Human Environment (the Stockholm Conference), 5–16 June 1972, Stockholm. Available online: <http://undocs.org/en/A/CONF.48/14/Rev.1> (accessed on 1 March 2023).
6. Report of the Brundtland Commission, 30 October 1987, United Kingdom. Available online: <https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf> (accessed on 2 March 2023).
7. Adoption of the Kyoto Protocol to the United Nations Framework Convention on Climate Change, 1 December 1997, Kyoto, Japan. Available online: <https://unfccc.int/decisions> (accessed on 4 March 2023).
8. Report of the Paris Agreement, 12 December 2015, Paris, France. Available online: <https://unfccc.int/process-and-meetings/the-paris-agreement> (accessed on 4 March 2023).
9. Declaration of Intention of Turkey according to the Paris Agreement. Available online: <https://www.mfa.gov.tr/paris-anlasmasi.tr.mfa> (accessed on 1 March 2023).
10. Lise, W. Decomposition of CO₂ emissions over 1980–2003 in Turkey. *Energy Policy* **2006**, *34*, 1841–1852. [\[CrossRef\]](#)
11. Hatzigeorgiou, E.; Polatidis, H.; Haralambopoulos, D. Energy CO₂ Emissions for 1990–2020: A decomposition analysis for EU-25 and Greece. *Energy Sources Part A* **2010**, *32*, 1908–1917. [\[CrossRef\]](#)
12. Carson, R.T.; Jeon, Y.; McCubbin, D.R. The relationship between air pollution emissions and income: US Data. *Environ. Dev. Econ.* **1997**, *2*, 433–450. [\[CrossRef\]](#)
13. Zhang, M.; Mu, H.; Ning, Y. Accounting for energy-related CO₂ emission in China 1991–2006. *Energy Policy* **2009**, *37*, 767–773. [\[CrossRef\]](#)
14. Gunduz, H.I. Review of relationship between environmental pollution and income: Panel cointegration analysis and error correction model. *Marmara Univ. J. Econ. Adm. Sci.* **2014**, *36*, 409–423.
15. Bruvold, A.; Medin, H. Factors behind the environmental Kuznets curve: A decomposition of the changes in air pollution. *Environ. Resour. Econ.* **2003**, *24*, 27–48. [\[CrossRef\]](#)
16. Deacon, R.T.; Norman, C.S. Does the environmental Kuznets curve describe how individual countries behave? *Land Econ.* **2006**, *82*, 291–315. [\[CrossRef\]](#)
17. Burnett, J.W.; Bergstrom, J.C.; Wetzstein, M.E. Carbondioxide emissions and economic growth in the US. *J. Policy Model.* **2013**, *35*, 1014–1028. [\[CrossRef\]](#)

18. Ang, C.B. Economic development, pollutant emissions and energy consumption in Malaysia. *J. Policy Model.* **2008**, *30*, 271–278. [\[CrossRef\]](#)
19. Chang, C.C. A multivariate causality test of carbon dioxide emissions, energy consumption and economic growth in China. *Appl. Energy* **2010**, *87*, 3533–3537. [\[CrossRef\]](#)
20. Magazzino, C. The relationship between CO₂ emissions, energy consumption and economic growth in Italy. *Int. J. Sustain. Energy* **2016**, *35*, 844–857. [\[CrossRef\]](#)
21. Mbarek, M.B.; Saidi, K.; Rahman, M.M. Renewable and non-renewable energy consumption, environmental degradation and economic growth in Tunisia. *Qual. Quant.* **2018**, *52*, 1105–1119. [\[CrossRef\]](#)
22. Sarkodie, S.A.; Strezov, V. Effect of foreign direct investments, economic development and energy consumption on greenhouse gas emissions in developing countries. *Sci. Total Environ.* **2019**, *646*, 862–871. [\[CrossRef\]](#)
23. Linh, D.H.; Lin, S.-M. CO₂ emissions, energy consumption, economic growth and FDI in Vietnam. *Manag. Glob. Transit.* **2014**, *12*, 219–232.
24. Grimes, P.; Jeffrey, K. Exporting the greenhouse: Foreign capital penetration and CO₂ emissions 1980–1996. *J. World Syst. Res.* **2003**, *2*, 261–275. [\[CrossRef\]](#)
25. Jorgenson, A.K. The effects of primary sector foreign investment on carbon dioxide emissions from agriculture production in less-developed countries 1980–99. *Int. J. Comp. Sociol.* **2007**, *48*, 29–42. [\[CrossRef\]](#)
26. Yang, W.-P.; Yang, Y.; Xu, J. The impact of foreign trade and FDI on environmental pollution. *China-USA Bus. Rev.* **2008**, *7*, 1–11.
27. Ahmed, K.; Shahbaz, M.; Qasim, A.; Long, W. The linkages between deforestation, energy and growth for environmental degradation in Pakistan. *Ecol. Indic.* **2015**, *49*, 95–103. [\[CrossRef\]](#)
28. Al-Mulali, U.; Weng-Wai, C.; Sheau-Ting, L.; Mohammed, A.H. Investigating the environmental Kuznets curve (EKC) hypothesis by utilizing the ecological footprint as an indicator of environmental degradation. *Ecol. Indic.* **2015**, *48*, 315–323. [\[CrossRef\]](#)
29. Apergis, N.; Ozturk, I. Testing environmental Kuznets curve hypothesis in Asian countries. *Ecol. Indic.* **2015**, *52*, 16–22. [\[CrossRef\]](#)
30. Lau, L.S.; Choong, C.K.; Eng, Y.K. Investigation of the environmental Kuznets curve for carbon emissions in Malaysia: Do foreign direct investment and trade matter? *Energy Policy* **2014**, *68*, 490–497. [\[CrossRef\]](#)
31. Jalil, A.; Mahmud, S.F. Environment Kuznets curve for CO₂ emissions: A cointegration analysis for China. *Energy Policy* **2009**, *37*, 5167–5172. [\[CrossRef\]](#)
32. Shahbaz, M.; Khraief, N.; Uddin, G.S.; Ozturk, I. Environmental Kuznets curve in an open economy: A bounds testing and causality analysis for Tunisia. *Renew. Sustain. Energy Rev.* **2014**, *34*, 325–336. [\[CrossRef\]](#)
33. Shahbaz, M.; Dube, S.; Ozturk, I.; Jalil, A. Testing the environmental Kuznets curve hypothesis in Portugal. *Int. J. Energy Econ. Policy* **2015**, *5*, 475–481.
34. Menyah, K.; Rufael, Y.W. CO₂ emissions, nuclear energy, renewable energy and economic growth in the US. *Energy Policy* **2010**, *38*, 2911–2915. [\[CrossRef\]](#)
35. Narayan, P.K.; Narayan, S. Carbon dioxide emissions and economic growth: Panel data evidence from developing countries. *Energy Policy* **2010**, *38*, 661–666. [\[CrossRef\]](#)
36. Apergis, N.; Payne, J.E. Renewable energy consumption and growth in Eurasia. *Energy Econ.* **2010**, *32*, 1392–1397. [\[CrossRef\]](#)
37. Tugcu, C.T.; Ozturk, I.; Aslan, A. Renewable and non-renewable energy consumption and economic growth relationship revisited: Evidence from G7 countries. *Energy Econ.* **2012**, *34*, 1942–1950. [\[CrossRef\]](#)
38. Leita, N.C. Economic growth, carbon dioxide emissions, renewable energy and globalization. *Int. J. Energy Econ. Policy* **2014**, *4*, 391–399.
39. Apergis, N.; Danuletiu, D.C. Renewable energy and economic growth: Evidence from the sign of panel long-run causality. *Int. J. Energy Econ. Policy* **2014**, *4*, 578–587.
40. Boluk, G.; Mert, M. Fossil & renewable energy consumption, GHGs (greenhouse gases) and economic growth: Evidence from a panel of EU (European Union) countries. *Energy* **2014**, *74*, 439–446.
41. Ang, J.B. CO₂ emissions, energy consumption, and output in France. *Energy Policy* **2007**, *35*, 4772–4778. [\[CrossRef\]](#)
42. Kanjilal, K.; Ghosh, S. Environmental Kuznet's curve for India: Evidence from tests for cointegration with unknown structural breaks. *Energy Policy* **2013**, *56*, 509–515. [\[CrossRef\]](#)
43. Hamzacebi, C.; Karakurt, I. Forecasting the energy-related CO₂ emissions of Turkey using a grey prediction model. *Energy Sources Part A Recovery Util. Environ. Eff.* **2015**, *37*, 1023–1031.
44. Halicioglu, F. An econometric study of CO₂ emissions, energy consumption, income and foreign trade in Turkey. *Energy Policy* **2009**, *37*, 1156–1164. [\[CrossRef\]](#)
45. Mutafoğlu, T.H. Foreign direct investment, pollution, and economic growth evidence from Turkey. *J. Dev. Soc.* **2012**, *28*, 281–297.
46. Polat, M.A. The analysis of the relationship between foreign capital investments and CO₂ emissions in Turkey through structural break tests. *J. Int. Soc. Res.* **2015**, *8*, 1127–1135.
47. Balibey, M. Relationships among CO₂ Emissions, economic growth and foreign direct investment and the environmental Kuznets curve hypothesis in Turkey. *Int. J. Energy Econ. Policy* **2015**, *5*, 1042–1049.
48. Kizilkaya, O. The impact of economic growth and foreign direct investment on CO₂ emissions: The case of Turkey. *Turk. Econ. Rev.* **2017**, *4*, 106–118.

49. Saatci, M.; Dumrul, Y. The relationship between economic growth and environmental pollution: The estimation of environmental Kuznets curve with a cointegration analysis of structural breaks for Turkish economy. *Erciyes Univ. J. Econ. Adm. Sci.* **2011**, *37*, 65–86.
50. Yavuz, N.C. CO₂ emission, energy consumption, and economic growth for Turkey: Evidence from a cointegration test with a structural break. *Energy Sources Part B Econ. Plan. Policy* **2014**, *9*, 229–235. [[CrossRef](#)]
51. Ozturk, I.; Acaravci, A. The long-run and causal analysis of energy, growth, openness and financial development on carbon emissions in Turkey. *Energy Econ.* **2013**, *36*, 262–267. [[CrossRef](#)]
52. Shahbaz, M.; Ozturk, I.; Afza, T.; Ali, A. Revisiting the environmental Kuznets curve in a global economy. *Renew. Sustain. Energy Rev.* **2013**, *25*, 494–502. [[CrossRef](#)]
53. Ocal, O.; Aslan, A. Renewable energy consumption-economic growth nexus in Turkey. *Renew. Sustain. Energy Rev.* **2013**, *28*, 494–499. [[CrossRef](#)]
54. Altintas, H. The relationships between primary energy consumption, carbon dioxide emission and economic growth in Turkey: An analysis of cointegration and causality. *Eskiseh. Osman. Univ. J. Econ. Adm. Sci.* **2013**, *8*, 263–294.
55. The Turkey Page of the Worldbank Website. Available online: <https://data.worldbank.org/country/Turkiye> (accessed on 1 March 2023).
56. Cleveland, R.B.; Cleveland, W.S.; McRae, J.E.; Terpenning, I. STL: A seasonal-trend decomposition procedure based on loess. *J. Off. Stat.* **1990**, *6*, 3–73.
57. Dickey, D.A.; Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **1979**, *74*, 427–431.
58. Phillips, P.C.B.; Perron, P. Testing for a unit root in time series regression. *Biometrika* **1988**, *75*, 335–346. [[CrossRef](#)]
59. Kwiatkowski, D.; Phillips, P.C.B.; Schmidt, P.; Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econom.* **1992**, *54*, 159–178. [[CrossRef](#)]
60. Yucesan, M. Sales forecast with YSA, ARIMA and ARIMAX methods: An application in the white goods sector. *J. Bus. Res.* **2018**, *10*, 689–706.
61. Bierens, H.J. ARMAX model specification testing, with an application to unemployment in the Netherlands. *J. Econom.* **1987**, *35*, 161–190. [[CrossRef](#)]
62. Arya, P.; Paul, R.K.; Kumar, A.; Singh, K.N.; Sivaramne, N.; Chaudhary, P. Predicting pest population using weather variables: An ARIMAX time series framework. *Int. J. Agric. Stat. Sci.* **2015**, *11*, 381–386.
63. Sutthichaimethee, P.; Ariyasajakorn, D. Forecasting energy consumption in short-term and long-term period by using ARIMAX model in the construction and materials sector in Thailand. *J. Ecol. Eng.* **2017**, *18*, 52–59. [[CrossRef](#)]
64. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
65. Murtagh, F. Multilayer perceptrons for classification and regression. *Neurocomputing* **1991**, *2*, 183–197. [[CrossRef](#)]
66. van der Walt, S.; Colbert, S.C.; Varoquaux, G. The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.* **2011**, *13*, 22–30. [[CrossRef](#)]
67. Kramer, O. *SciKit-Learn in Machine Learning for Evolution Strategies*; Springer: Berlin/Heidelberg, Germany, 2018.
68. Lemenkova, P. Processing oceanographic data by python libraries NumPy, SciPy and Pandas. *Aquat. Res.* **2019**, *2*, 73–91. [[CrossRef](#)]
69. The Turkey Page of the Our World in Data Based on the Global Carbon Project Website. Available online: <https://ourworldindata.org/co2/country/turkey> (accessed on 1 March 2023).
70. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [[CrossRef](#)]
71. Mombeini, H.; Chamzini, A.Y. Modeling gold price via artificial neural network. *J. Econ. Bus. Manag.* **2015**, *3*, 699–703. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.