Applied Energy 328 (2022) 120163

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

An interpretable forecasting framework for energy consumption and CO_2 emissions

Serkan Aras [a,*], M. Hanifi Van [b]

[a] Econometrics Department, Faculty of Economics and Administrative Sciences, Dokuz Eylul University, Izmir, Turkey
[b] Econometrics Department, Faculty of Economics and Administrative Sciences, Van Yuzuncu Yil University, Van, Turkey

HIGHLIGHTS

- An interpretable forecasting framework is proposed.
- The drivers of total energy consumption and CO_2 emissions in Turkey are identified.
- A comprehensive comparison between different machine learning methodologies is carried out.
- The interactions of machine learning models with features selection techniques are investigated.

ARTICLE INFO

Keywords:
Machine learning
Model Interpretability
SHAP
Energy consumption
CO_2 emissions

ABSTRACT

It is a well-established fact that energy consumption and production, as the primary sources of greenhouse gases, contribute to climate change and global warming issues. The analysis and estimation of the factors that contribute to these harmful gases will be of great assistance in the development of policies to reduce carbon dioxide emissions. In addition to identifying the factors related to energy consumption and CO_2 emissions, forecasting the variable of interest as accurately as possible has a key role in increasing the efficiency of energy strategies to be implemented. Unlike studies in the literature, this study not only forecasts the future value of energy consumption and CO_2 emissions but also determines the relationship between the predictions and the influential variables by revealing the contribution of each variable to the prediction. For this purpose, the study proposes an interpretable forecasting framework based on values of the Shapley additive explanation (SHAP) to provide a simpler explanation of machine learning (ML) models in forecasting energy consumption and CO_2 emissions. The results obtained show that the total electricity generation from different energy sources is found to be the most important variable interacting positively with both energy consumption and CO_2 emissions. Also, the influence of the predictors on projections made before and after COVID-19 has changed dramatically. The proposed method may assist policymakers in making future energy investments and establishing energy laws more accurately and efficiently as it explains the drivers of the forecasts.

1. Introduction

There has been a steady rise in the need for energy since the industrial revolution began. In addition, the socio-economic growth of societies has changed the vital activities of residents, which has increased the need for energy even more. Today's society would be unable to function without energy. A country's degree of development and industrialisation correlate directly with its energy consumption or demand [1]. These advancements hastened the industrialization of countries and facilitated the emergence of economic growth and development in countries that could quickly adapt to this process and closely follow technological advancements [2]. Countries' increased output for growth and development has necessitated a rise in energy consumption [3,4]. Energy consumption and environmental pollution have become inextricably linked since the industrial revolution. In developed economies and developing ones, such as Turkey, the growth in greenhouse gas emissions, especially after 1990, has generated worries about global warming. Under the direction of the United Nations, this situation has led to the adoption at the global level of different regulations and measures concerning greenhouse gases. In this context,

* Corresponding author.
E-mail address: serkan.aras@deu.edu.tr (S. Aras).

https://doi.org/10.1016/j.apenergy.2022.120163
Received 10 June 2022; Received in revised form 22 September 2022; Accepted 14 October 2022
Available online 22 October 2022

it is crucial for decision-makers to assess the pollution produced by energy consumption, and then design an energy policy that includes preventative steps to combat energy pollution.

All of this, together with the relationship between energy use, CO_2 emissions, and economic growth, has significantly increased public concern about climate change. Global warming is mostly caused by human-caused emissions of greenhouse gases, particularly carbon dioxide. Climate change researchers are increasingly concerned about reducing CO_2 emissions as a means of reducing greenhouse gas emissions [5]. Climate change and environmental degradation have been compounded by excessive carbon dioxide emissions, which cause a major danger to global security and human well-being. Changes to the amount and patterns of precipitation will occur as temperatures rise, and so will an increase in extreme weather [6,7].

For many researchers, global warming is linked to the rapid expansion of the global economy as well as to greenhouse gas emissions that affect the Earth's climate [8]. With global greenhouse gas emissions increasing, the United Nations signed the Kyoto Protocol in 1997 and join the agreement in 2005, imposing enforceable responsibilities on developed countries. In the first period of the protocol, 191 countries committed to reducing their emissions by 5 per cent between 2008 and 2012 compared to 1990. Although it was resolved to reduce emissions during the second commitment period, by at least 18 per cent in 2020 the resolution has not entered into force because not enough countries signed it. Another important agreement, the Paris climate agreement, was signed in 2015 to help people and businesses be better able to deal with climate change after 2020. With this agreement, 196 countries pledged to reduce greenhouse gas emissions, constraining the global temperature rise to less than 2 degrees Celsius until 2030. Turkey signed the Kyoto Protocol in 2009 and the Paris climate agreement in 2016. Turkey is the 37th largest country in the world in terms of land size, it is located on the Asian and European continents, and has a population of approximately 83 million as of 2020. It is a member of the G20, a group of the 20 countries with the world's largest economies and whose energy consumption has increased significantly.

According to World Bank data (2020), Turkey has the world's 19th largest economy, with a nominal GDP of $720 million [9]. Turkey's energy consumption has grown across all sectors since 2001, except for industry, where it decreased slightly in 2018. Between 2008 and 2018, Turkey's energy consumption increased by 86 per cent in transportation, 60 per cent in industry, and 12 per cent and 2 per cent in services and residential sectors respectively [10]. The greenhouse gas inventory data indicate that the overall greenhouse gas emissions for 2020 would be 523.9 million tons (MT) of CO_2, increasing 3.1 per cent from the previous year. Total greenhouse gas emissions per capita were expected to be 4 tons of CO_2 in 1990, 6.2 tons in 2019, and 6.3 tons in 2020 [11]. Given the importance of the subject, this study uses ML techniques to forecast the amount of CO_2 emissions and energy consumption in Turkey.

The black-box nature of machine learning methods is one of the biggest drawbacks of these techniques, preventing them from being widely used in problems that require transparency and involve high risk and cost. Over the last five years, many innovative methods have been developed to tackle this problem [12–16]. The core idea behind these methods is to generate more interpretable and explainable ML models by revealing the contribution of each covariate to the final prediction. These methods are generally classified as model-specific and model-agnostic interpretations, of which the former works with model internal interpretations and is designed for a specific model whereas the latter works with any ML model and relies on examining the behaviour of the input and output pairs. Given the scope of this study, we preferred a model-agnostic explanation since it provides a general approach, and we tested various ML models, each of which has different inner working mechanisms. Among model-agnostic methods, SHapley Additive exPlanations (SHAP), proposed by Lundberg et al. [15], is different from the others due to appealing features such as consistency, local accuracy, and missingness. That is the reason why the SHAP method is employed here in combination with ML models to explain the predictions.

The contributions of this paper are threefold. The first contribution is to expand the literature in the field of energy by acting as a bridge between two different research methodologies. The studies in the literature focus on either finding a possible relationship between CO_2 emissions (or energy consumption) and some specific variables [17–20] or forecasting the future values of these critical variables to control them by making policies associated with them [1,5,21–23]. The contribution of this study is to fill the gap between these two fields by investigating the best forecasting model for CO_2 emissions and energy consumption and by identifying the possible relationship between the predictors and predictions both locally and globally by finding the contribution value of every feature to the prediction value. Furthermore, the majority of the studies have made long-term predictions for energy consumption/demand and CO_2 emissions [21,24,25]. Even if the long-term forecasts are still of value for policymakers to take measures to reduce the carbon footprint and increase the share of renewable energy within the energy consumption spectrum, it can be tricky because of possible and unpredictable radical changes in the long term, such as climate change and technological developments in energy generation. Hence, this study adds new evidence to the literature on short-term forecasting.

The second contribution is to propose an interpretable forecasting framework, which is shown in Fig. 1, based on ML models. This framework combines feature selection techniques with ML models to obtain more accurate forecasts and a more reliable explanation of the problem under investigation. Benefitting from feature selection has two functions here. First, to eliminate irrelevant or redundant variables from the analysis to increase the forecasting performance. Second, to avoid the tendency of SHAP methods to assign Shapley values to the variables which have low or no prediction power. In this regard, the feature selection technique performs a filtering task in deciding which variables will have a Shapley value to explain the prediction and which ones will have a Shapley value of zero. The intention is to obtain a simple explanation of the problem. For this purpose, we employed three feature selection techniques: Boruta, orthogonal matching pursuit (OMP), and least absolute shrinkage and selection operator (LASSO); and six ML models: random forest (RF), artificial neural networks (ANN), support vector machines (SVM), k-nearest neighbours (KNN), Adaboost, and gradient boosting decision trees (GBDT).

The final of this paper is to provide new evidence to the scant literature available on SHAP explanations to ML models. Although the SHAP method has recently found applications in various fields, more evidence is still needed to demonstrate its effectiveness in providing insight into how black box models produce their predictions. It is worth mentioning some studies using ML models combined with the SHAP method in different application areas: solar power forecasting [26], demand modelling [27], prediction of natural ventilation rate, fraud detection, real-time accident detection [28], gold price forecasting [29], electricity price forecasting, energy consumption in electric vehicles [30], prediction of COVID-19 diagnosis [31], and customer purchase forecasting [32]. To the best of our knowledge, this study is the first that applies the SHAP method to the analysis of total energy consumption and CO_2 emissions to find out the drivers that influence the predictive performance.

2. Related literature

There is a strong link between economic growth and energy use since the need for energy will go up as the economy grows [33]. The link between energy consumption, environmental pollutants, and economic growth has been discussed as three research standards in the literature [34]. The first standard is studies that examine the relationship between environmental pollution and economic growth within the context of the Environmental Kuznets Curve (EKC) hypothesis [35]. The EKC hypothesis, introduced by [36,37], states that environmental pollution will

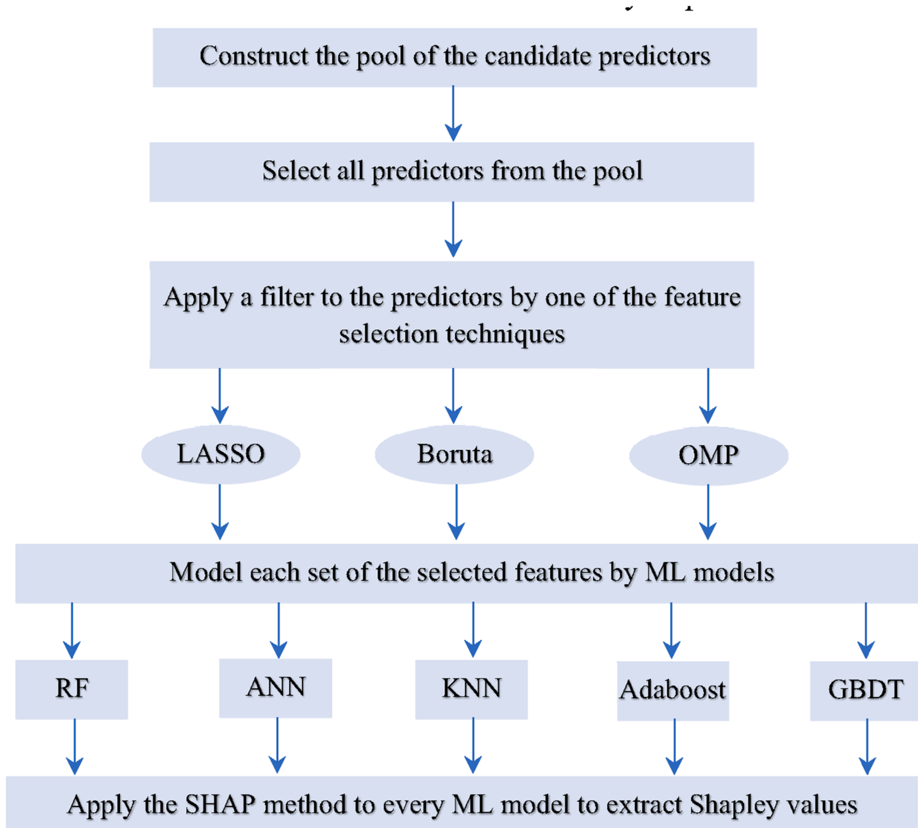

Fig. 1. The demonstration of the proposed method.

decrease after a certain threshold value of growth [18,38]. The second standard focuses on the link between income and energy consumption. The third standard is a combination of the two approaches above and deals with the dynamic link between energy consumption and environmental pollutants [34,35].

Numerous studies show that carbon dioxide emissions are linked to GDP, energy consumption, and vehicle use. These studies use a variety of modelling methodologies and cover different periods and countries. The research is typically classified into two categories, i.e. structural analyses and predictions, and analysis methodologies are employed accordingly. Although econometric approaches are more prevalent in structural analysis, machine learning is more prevalent in predicting. Transportation [39,40], industrialisation [7,17,41], economic growth, and renewable energy [42,43] have typically been the focus of these types of studies. Table 1 provides a summary of studies on energy use and CO_2.

3. Machine learning techniques

Brief information about machine learning and variable selection techniques used in the analysis is given in this and the following section. In addition, Table 2 provides the definitions of the parameters in the study.

3.1. Rf

Random forest (RF) is based on the idea of ensemble learning. It has multiple-decision classification trees, it randomly picks features from each tree, and then it votes or averages the results. The results are stable and accurate predictions. Assume there is an "input" dataset, and one decision tree predicts that it will be $\{H(x,\theta_i)\}$. The random forest prediction model will use all these trees' predictions to come up with its final result, which will be the average of the predictions made by each one [48]. The following equation describes the prediction made by a RF model:

$$\overline{H}(x) = \frac{1}{x}\sum_{i=1}^{k} H(x,\theta_i) \tag{1}$$

where $\overline{H}(x)$ is a predicted value, θ_i is the random variable of a decision tree.

3.2. ANN

ANNs were invented in 1943 by neurophysiologist Warren McCulloch and mathematician Walter Pitts [49]. ANNs are a class of models that are useful for tackling issues involving function approximation, pattern recognition, classification, and clustering. Neural networks are generated from many small computing units named artificial neurons. An artificial neuron, like a biological neuron, contains a body in which computations are done, as well as some input channels and one output channel [50]. Every node in a neural network contains an output function, and it has three layers: an input layer, a hidden layer, and a layer that outputs information. The connection between the two neurons is assigned a weight during the training phase and is tested on test datasets [51]. The primary disadvantage of the ANN technique is that it prevents the researcher from examining the direct link between the input and output variables [45].

The advantage of ANNs over other models is their capacity to simulate multivariate problems characterized by complicated connections between variables. The non-linear correlations between these variables may also be extracted by ANN through the process of learning training data [52]. It is worth noting that multi-layer perceptron is the type of ANN model utilised in this study.

Table 1
Summary of previous literature.

Author	Period	Region	Method	Variables	Key findings
[44]	1978–2012	G7	FMOLS, DOLS	EC, Y, EP, GP, RP	Home electricity demand is price-elastic but income-inelastic over the long run. Further, climate regulations in the G7 countries have the potential to lower domestic power demand and, consequently, carbon emissions over time.
[17]	1995–2018	23 developing countries	ARDL, Heterogeneous causality	CO_2, GDP, EU, INDUS, URBAN	CO_2 emissions are influenced by GDP growth, energy use, industrialisation, and urbanisation.
[18]	1975–2000	China	STIRPAT	CO_2, EI, GDP, Population,	CO_2 emissions are affected by factors such as population, prosperity, and technology.
[19]	1970–2013	69 selected countries	VECM, FMOLS, DOLS	CO_2, GDP, RE, Energy use,	Long-term causal links exist between CO_2, GDP, and energy use.
[20]	1971–2015	China	ARDL	CO_2, EC, GDP, RE, NRE, EX, IM, Eİ	In the short term, the participation of the private sector in energy investment can raise carbon emissions, but in the long term, it can reduce carbon emissions.
[21]	1970–2016	Turkey	DL, ANN, SVM	CO_2, Transportation energy demand, GDP, Population, Vehicle	Turkey's yearly growth rates for transportation-related energy consumption and CO_2 emissions will increase by 3.7 percent and 3.65 percent, respectively.
[22]	1971–2018	Turkey	ANN, LSTM	EC, CO_2, GDP	They discover that when nations are classified according to income level, those with middle-low and poor incomes expand more rapidly under moderate growth policies in terms of carbon emissions and energy consumption, validating the EKC hypothesis.
[45]	1989–2008	Thailand	ANN, OLS	Energy demand, GDP, Population, vehicle	Thailand's transportation energy consumption will rise to 54.4–59.1 million tonnes of oil equivalent in 2030, or around 237–256 percent of 2008 levels.
[1]	1975–2016	Turkey	ANN	Energy demand GDP, population, Ton-km, Vehicle, Oil price	The study concludes that the model with the lowest error rates and the highest R2 was the best performer. It considered the price of oil, the population, and kilometres covered by motor vehicles.
[40]	1970–2001	Turkey	ANN	Transport energy demand, GNP, Population, Vehicle	The ANN models capture the fluctuation in both dependent and independent variables in historical data.
[46]	2005–2018	China	GM, V-GM	CO_2, cement industry	In terms of simulating the actual CO_2 emissions of the cement sector, V-GM (1, N) has the highest accuracy at 97.29 percent.
[23]	2013–2018	Turkey	ANN-TLBO, ANN-BP, ANN-ABC	Electric energy demand, GDP, I, E, GEED, Population,	ANN-TLBO models outperform ANN-BP and ANN-ABC models in EED estimation. The ANN-TLBO model reduced the average root-mean-square error of the ANN-BP and the ANN-ABC models by 42.3 percent and 39.3 percent, respectively.
[47]	2000–2025	Iran	SDM	EC, CO_2, ES, Energy production,	Total CO_2 emissions will reach 985 million tonnes in 2025, representing an annual increase of around 5 percent.
[5]	1978–2012	China	LSSVM, BPNN, GM	CO_2, GDP, EC	The classification of total CO_2 emissions helps improve the accuracy of predicting CO_2 emissions.

Table 2
The nomenclature used in the study.

Symbol	Meaning
x	The characteristic variable
k	The number of decision trees
θ_i	The random variable of a decision tree
$H(x, \theta_i)$	The predicted value by a decision tree
$\overline{H}(x)$	The predicted value by a random forest
T	The size of the sample
M	The number of features
λ	The hyperparameter that determines the amount of shrinkage
$\phi^h_{t,j}$	The Shapley value of the j feature at time t predicted for horizon h
$\phi^h_{t,0}$	The average prediction value when no feature contributes to the prediction.
S	The features in a coalition
$f(S)$	The prediction made by the coalition
$\widehat{\beta}_h$	The coefficient vector of the LASSO model for horizon h
y_i	The true value of the i th observation
$\widehat{y}_i$	The forecast value of the i th observation

3.3. KNN algorithm

The KNN method derives its name from the fact that it classifies unlabelled samples by utilizing information about an example's k-nearest neighbours. Any number of nearest neighbours can be represented by the letter k. After determining k, the method requires a training dataset comprised of instances categorized into many categories using a nominal variable. For each unlabelled record in the test dataset, KNN selects the k records in the training data that are the "nearest" in similarity. It assigns an unlabelled test instance to the class that most closely corresponds to its k-nearest neighbours [53].

3.4. SVM

SVM is a popular machine learning technique derived from statistical learning theory. It started as a mechanism to help with sample classification and evolved into something more useful: a forecasting tool [54]. SVM's training is similar to solving a linearly restricted quadratic programming problem; therefore the solution of SVM is always unique and globally optimal, unlike neural networks training, which involves non-linear optimization with the risk of becoming trapped in local minima [55].

3.5. Adaboost

The Adaboost algorithm was first put forward by Freund and Schapire [56] to improve the performance of a weak learner by reducing its variance and bias together. There are many versions of the Adaboost algorithm in the literature. This study uses the variant developed by Drucker [57]. The core idea behind the algorithm is to sequentially update the predictions of the previous weak learner by giving higher weight values to the bigger errors. Every weak learner tries to predict and minimise the errors of the previous learner. Hence, the final prediction of the algorithm is obtained by summing all predictions over the iterations. Although decision trees have been traditionally employed as a base learner, any forecasting method or model can be exploited. Learning rate and the number of base learners are two important parameters of the algorithm that need to be fine-tuned carefully.

3.6. GBDT

GBDT was proposed by Friedman [58] as an improved version of the Adaboost algorithm. GBDT relies on the idea of improving predictions gradually within a functional gradient descent framework. It starts with its first prediction by taking an average of the variable to be predicted and continues the algorithm by adding the prediction of decision trees to the previous fitting values until a stopping criterion is met. Unlike Adaboost, GBDT employs bigger decision trees as a weak learner. This weak learner can be configured by using tree hyperparameters and exploits a loss function, which can be any form as long as it is differentiable, to minimise errors by the gradient descent method. Like Adaboost, the final prediction is obtained by summing all predictions in the iterations.

3.7. Shapley values

Model interpretability approaches can be divided into two categories, i.e. local and global. The local methods correspond to the evaluation of a single prediction, whereas global interpretability allows the assessment of the general relationship between a feature and the output. Fortunately, the SHAP method provides both a local and a global explanation. The SHAP explanation is based on the Shapley values which are calculated in cooperative game theory to determine how the total payoff will be shared among players in a coalition according to their contribution [59]. The study by Štrumbelj and Kononenko [60] was the first to make an analogy between the players of a game and the covariates of a model. Following this study, Lundberg et al. [15] put the Shapley values into the framework of additive feature attribution (SHAP), which is a linear model, and proposed two special estimation methods, KernelSHAP and TreeSHAP.

Suppose that the prediction at time t made for horizon h is represented by $\widehat{y}_{t+h}$. It is possible to express this prediction as the sum of the Shapley values as follows by Eq. (2):

$$\widehat{y}_{t+h} = \phi^h_{t,0} + \sum_{j=1}^{M} \phi^h_{t,j} \equiv \Phi^h_t \tag{2}$$

where $\phi^h_{t,j}$ corresponds to the Shapley value of the j th feature and M represents the number of features. $\phi^h_{t,0}$ is the base value showing the average prediction value when no feature contributes to the prediction.

To find a contribution of a feature precisely, all possible subsets of features (coalitions) must be evaluated. This results in the calculation of the average marginal contribution as given as follows by Eq. (3):

$$\phi^h_{t,j} = \sum_{S\subseteq M\setminus j} \frac{|S|!(M-|S|-1)!}{M!}[f(S\cup\{j\}) - f(S)] \tag{3}$$

where S is the features in the coalition and $f(S)$ denotes the prediction made by this coalition.

4. Feature selection techniques

4.1. LASSO

The LASSO is a penalised regression model which is devised by Tibshirani [61] to tackle the deficiency of ridge regression in producing small but not zero coefficients. The LASSO changes the penalty function of ridge regression with L_1 norm defined as $\lambda\sum_{i=1}^{M}|\beta_{h,i}|$. This leads to the following estimator defined by Eq. (4) to find the coefficients of the regression model.

$$\widehat{\boldsymbol{\beta}}_h = argmin_{\beta_h}\left[\sum_{t=1}^{T-h}\left(y_{t+h} - \boldsymbol{\beta}_h^{'}\boldsymbol{x}_t\right)^2 + \lambda\sum_{i=1}^{M}|\beta_{h,i}|\right] \tag{4}$$

where T is the size of the sample and M represents the number of features while λ is a hyperparameter that determines the amount of shrinkage.

The coefficients equal to nonzero values are the selected features by the model. The LASSO even works when the number of variables is bigger than the sample size.

4.2. OMP

The OMP method [62] originates from the area of signal processing as a more advanced variant of the matching pursuit algorithm [63]. A matching pursuit algorithm is a greedy approach for solving the sparse optimisation issue. The OMP picks iteratively-one atom from training dictionary atoms to identify the atom that is most similar to the signal while reducing the signal's approximation error. It selects the next atom by finding the highest correlation between the residuals of the previous iteration and the atoms of the current iteration. As a way of increasing the convergence, the OMP algorithm utilises Schmitt orthogonalization to make the residuals orthogonal to the atoms that were already chosen.

4.3. Boruta algorithm

The Boruta is a feature selection technique, proposed by Kursa et al. [64], which includes a random forest model as a wrapper tool. Unlike other feature selection techniques, it tries to identify all relevant features by a statistical test. To remove the correlation between a feature and model output, the Boruta algorithm inserts a randomised copy of that feature, which is called the shadow feature, into a dataset. Then it compares the importance values assigned to the shadow feature and the corresponding feature and records the number of times a feature's importance value is more than the maximum Z score of shadow features. These numbers are then utilised to determine whether features are statistically important.

5. Dataset and empirical methodology

In this paper, we collected a set of variables considered highly related to CO_2 emissions and total energy consumption in Turkey to investigate the effects of these variables on forecasting performance. The data definitions and their sources are presented in Table 3. In addition, the descriptive statistics of the variables in the analysis for the full sample are given in Table 4. All models in this study are fitted two times. First, CO_2 emissions are taken as a dependent variable and the rest are the independent variables. Then, total energy consumption is taken as the dependent variable, and all analyses are repeated. The data cover the period from 1970 to 2021. The test data starting in 2001 with 20 yearly observations are forecasted by 24 machine learning models. This study collects annual data from government agencies and international organizations. The CO_2 emission and energy consumption data, which we used as dependent variables, were taken from the Turkish Statistical Institute (Turkstat, 2020) and the International Energy Agency [65], respectively. The data on electricity consumption, industrial electricity consumption, and the number of road motor vehicles were gathered from the Turkish Statistical Institute [66]. The data on oil consumption and energy supply were collected from the International Energy Agency [65], and the data on the gross domestic product (GDP) and population were gathered from the World Development Index [67].

Fig. 2 shows the correlation coefficient values between the variables. This figure consists of two parts. The upper part is based on the original values while the lower part is formed after making the series stationary, which is detected by the KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test, by applying a differencing operation. As you can see in the upper part,

Table 3
Definitions of variables.

Data	Definition	Source
CO_2 Emissions	Greenhouse gas emissions (Carbon dioxide equivalent)	TurkStat (Turkish Statistical Institute)
Oil: Consumption	Oil Consumption	International Energy Agency, (2022)
GDP	The annual growth of real GDP per capita	World Development Index 2020
Population	Population (millions)	World Development Index 2020
EGSER	Electricity generation and shares by energy resources (EGSER consists of coal, liquid fuels, natural gas, hydro, renewable energy, and waste totals)	TETC, Electricity Generation - Transmission Statistics of Turkey
EC (total)	Net electricity consumption (total)	TEDC, Electricity Distribution and Consumption Statistics of Turkey
EC (industrial)	Industrial net electricity consumption	TEDC, Electricity Distribution and Consumption Statistics of Turkey
Vehicle	Number of road motor vehicles (total)	General Directorate of Public Security, Notaries Union of Turkey
Total energy supply	Total energy supply	International Energy Agency, (2022)
Energy consumption	Total final energy consumption	International Energy Agency, (2022)

Table 4
Descriptive statistics for the full sample.

	Min	Mean	Q_1	Median	Q_3	Max	Std.Dev.	Skew.	Kurt.
CO_2 Emissions	39.28	177.9	82.84	154.7	266.6	397.2	104.4	0.53	−0.90
Energy consumption	15,155	51,622	28,453	46,910	73,588	105,035	26,258	0.50	−0.96
Oil: Consumption	7.10	25.84	16.93	27.96	31.62	49.26	10.63	0.31	−0.53
GDP (billion USD)	272.2	954.3	445.8	783.9	1322.8	2334	596.0	0.88	−0.37
Population (mil.)	35.30	58.83	47	59.30	69.97	82.60	13.89	−0.04	−1.24
EGSER	8623	111,199	26,751	82,285	187,744	304,802	93,470	0.65	−0.94
EC (total)	7308	91,117	23,807	64,398	152,119	258,232	77,349	0.73	−0.79
EC (industrial)	44.90	56.11	47.60	56.00	64.47	67.30	8.16	0.00	−1.78
Vehicle	3.7E + 05	7.9E + 06	1.9E + 06	5.8E + 06	1.3E + 07	2.3E + 07	7.0E + 06	0.77	−0.71
Total energy supply	18,213	67,721	34,197	58,626	96,969	146,814	37,549	0.60	−0.81

the variables exhibit an extremely high positive correlation with each other except in the case of the EC (industrial). However, this fact can be misleading since almost all variables have a strong trend component. Hence, the lower part is included to indicate more reliable correlation values after removing the trend component. This shows that there is still valuable information between the variables to be forecasted and the predictors. However, it should be noted that even though correlation values may provide insight into the potential relationship between variables, it only gives information about the linear relationship, and these values do not take into consideration the effect of the interactions between features on the fitted model. For this reason, in this study, feature selection techniques are effectively used to find out the features with high predictive power in the fitted model. The second reason for using feature selection techniques is to obtain simpler explanations regarding the problem at hand. The SHAP method tends to assign Shapley values to almost every predictor even if the predictor has a very low prediction power [68]. Using a feature selection can eliminate these non-important or non-informative features before applying the SHAP method.

An expanding window strategy was carried out by refitting all models for every next observation of test data. This means that our training data are extended by one observation every time, and the models are constructed 20 times for each dependent variable in the analysis. One-step-ahead forecasts are produced for CO_2 emissions and energy consumption at time *t* by using the values of independent variables at time (t-1) and the single lagged value of the dependent variable.

Finding the best values of hyperparameters is crucial to generating ML models with a good generalisation ability, and considerably affects the prediction performance of ML models. For this purpose, we carried out 5-fold cross-validation to tune hyperparameters of feature selection techniques and ML models. The search space for hyperparameters is presented in Table 5.

6. Evaluation statistics

To assess the forecasting performance of the ML techniques in this study, four error measures are used: mean absolute error (MAE), root mean squared error (RMSE), the coefficient of determination (R^2), and symmetric mean absolute percentage error (sMAPE). The first two are scale-dependent error measures. They can be calculated by the Eqs. (5)–(8). A more detailed discussion about the evaluation statistics can be found in the study by Hyndman [69].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{7}$$

$$sMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{(y_i + \hat{y}_i)}*200 \tag{8}$$

where y_i represents the true value of the variable and $\hat{y}_i$ denotes the forecasted value.

7. Results and discussions

This section presents the forecasting results obtained and a discussion on the interpretability of the best-performing ones for the two forecasting experiments that use 24 forecasting methods each of which is the combination of six ML models and four feature selection approaches. First, it compares the forecasting performances related to two

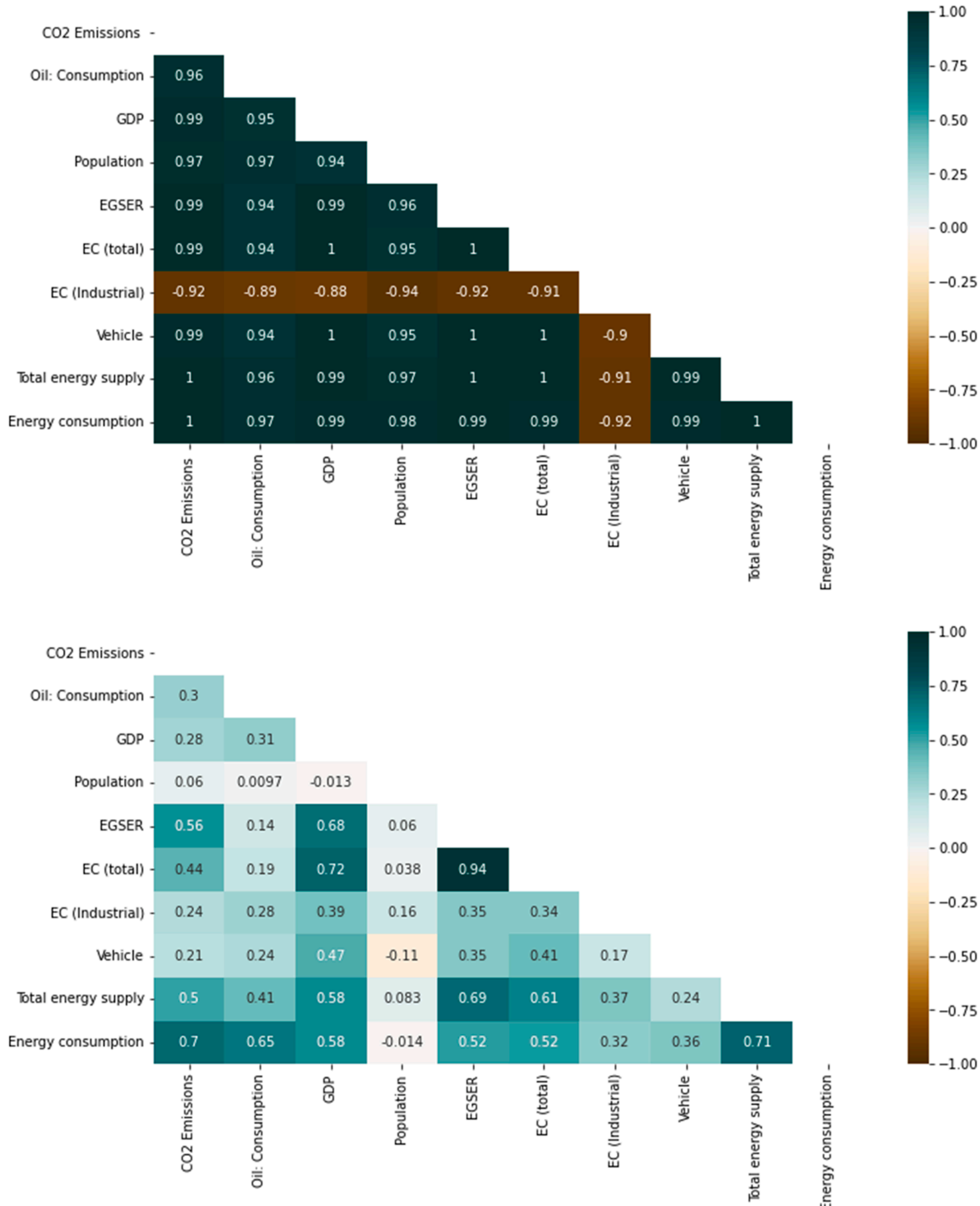

Fig. 2. Heatmaps for correlation coefficient values.

examined groups of data to identify which forecasting methodologies are superior to others in each case. Second, the discussion sheds some light on which predictors are more important than others in affecting the predictions of the best-performing models.

Table 6 includes the forecasting performances evaluated by four statistics for CO_2 emissions data. The results are grouped according to each ML model to easily show the effect of the different feature selection procedures on the investigated ML models. In each group, the superior model for every error measure is represented in bold. Furthermore, the model outperforming other models in all points of comparison is indicated with grey shading in the table. As can be seen in Table 6, feature selection techniques do not contribute to the performance improvements for RF and ANN models in forecasting CO_2 emissions. However, the other models enjoy a reduced number of features. For example, SVM and GBDT produce better forecasts with the features chosen by OMP while KNN works better with LASSO and Adaboost work better with the mix of LASSO and Boruta. When the overall performance is evaluated, ANN models dominate all the others across all the error measures considered, except for ANN with OMP. We have found that the best-performing two-model specifications are ANN and ANN with Boruta, respectively. In general, SVM is the worst model, leading to poor performance values compared to the other ML models.

As for the assessment of forecasting results for energy consumption data, Table 7 is constructed in the same way as Table 6. We can see that the results obtained in Table 6 and Table 7 are completely different. This may be an indication that the functional form and the relationship between variables differ considerably for CO_2 emissions and energy consumption in Turkey. We will examine this further with the help of the analysis of Shapley values. From Table 7, it is understood that feature elimination has a positive effect on achieving more accurate forecasts for

Table 5
The search space for hyperparameter tuning.

Model	Search Space
SVM	C ∈ [10000, 5000, 1000, 500, 100, 10, 1]
	γ ∈ [0.1, 0.01, 0.001, 0.0001, 0.00001]
	ε ∈ [0, 0.1, 0.01, 0.001]
	Kernel ∈ [Radial basis function]
RF	The number of trees ∈ [50, 100, 200, 300]
	The minimum samples to split a node ∈ [2,3,5,7,10]
	The minimum samples to be at a leaf node ∈ [1,3,5,7,10]
	The maximum depth of a tree ∈ [1,3,5,7,10]
KNN	The number of neighbours ∈ [1, 2, 3, …., 24, 25]
	Distance metric ∈ [Manhattan, Minkowski]
	Weight function ∈ [Uniform, Distance]
ANN	Batch size ∈ [2, 4, 8,]
	Alpha ∈ [1000, 100, 10, 1, 0.1, 0.01, 0.001, 0.0001]
	The size of hidden layer ∈ [(1,1), (5,1), (10,1), (30,1), (1,2), (5,2), (10,2), (30,2)]
	Maximum iterations ∈ [100, 200, 300]
	Activation function ∈ [the rectified linear unit function]
Adaboost	The number of estimators ∈ [15, 30, 50, 100, 200, 300, 500]
	Learning rate = [0.0001, 0.001, 0.005, 0.01, 0.03, 0.05, 0.07, 0.1, 0.15, 0.2, 0.3]
GBDT	The minimum samples to split a node ∈ [2,3,5,7,10,15,20]
	The minimum samples to be at a leaf node ∈ [1,3,5,7,10,15,20]
	The maximum depth of a tree ∈ [1,3,5,7,10]
	Learning rate ∈ [0.001, 0.005, 0.01, 0.03, 0.05, 0.07, 0.1, 0.2]
	The number of estimators ∈ [30, 50, 100, 200, 300]
	The loss function ∈ [absolute error]

Table 6
Forecasting results for CO_2 emissions data.

	MAE	RMSE	R^2	sMAPE
RF	**13.792**	**16.372**	**0.937**	**4.693**
RF with OMP	14.144	16.520	0.936	4.788
RF with Boruta	14.032	16.634	0.935	4.824
RF with LASSO	14.487	17.421	0.929	5.001
ANN	**11.513**	**14.062**	**0.954**	**3.974**
ANN with OMP	13.584	15.979	0.940	4.154
ANN with Boruta	11.911	14.758	0.949	4.740
ANN with LASSO	12.328	14.966	0.948	4.227
SVM	15.154	17.979	0.924	5.202
SVM with OMP	**14.462**	**16.905**	**0.933**	**5.116**
SVM with Boruta	19.811	31.759	0.764	6.947
SVM with LASSO	15.451	18.025	0.924	5.306
KNN	13.483	16.051	0.940	4.689
KNN with OMP	13.644	16.099	0.939	4.693
KNN with Boruta	13.757	16.531	0.936	4.812
KNN with LASSO	**13.398**	**15.954**	**0.940**	**4.640**
Adaboost	14.350	16.626	0.935	4.966
Adaboost with OMP	16.811	20.829	0.898	5.690
Adaboost with Boruta	**13.843**	16.426	0.937	**4.754**
Adaboost with LASSO	14.027	**16.169**	**0.939**	4.885
GBDT	13.025	15.543	0.943	4.472
GBDT with OMP	**12.660**	**15.125**	**0.946**	**4.345**
GBDT with Boruta	12.736	15.230	0.946	4.379
GBDT with LASSO	13.159	15.678	0.942	4.516

all ML models except KNN. RF and ANN enjoy OMP and Boruta selection techniques, Adaboost is better by using OMP and LASSO, and SVM works well with Boruta. Among ML models, the models based on a GBDT model result in the best-performing models in terms of all evaluation statistics. Specifically, GBDT with LASSO emerges as the model that produces the most accurate forecasts for energy consumption.

Now it is time to find out the drivers of the forecasts of the best-performing ML models detected by the forecast comparison made for CO_2 emissions and energy consumption data. For this purpose, the Shapley values produced by the SHAP method are used. Fig. 3 shows the bar plot obtained for each predictor of CO_2 emissions, where the x-axis represents the mean absolute Shapley values and the y-axis denotes the features sorted vertically according to the value importance in the x-axis. The colour of a bar indicates the direction of the general

Table 7
Forecasting results for energy consumption data.

	MAE	RMSE	R^2	sMAPE
RF	3606	4128	0.933	4.670
RF with OMP	3530	**4030**	**0.936**	4.519
RF with Boruta	**3402**	4107	0.934	**4.376**
RF with LASSO	3501	4082	0.935	4.567
ANN	3329	3858	0.942	4.410
ANN with OMP	3229	**3756**	**0.945**	**4.181**
ANN with Boruta	**3164**	3866	0.941	4.196
ANN with LASSO	3582	4035	0.936	4.723
SVM	3457	4198	0.931	4.449
SVM with OMP	3712	4200	0.931	4.925
SVM with Boruta	**3207**	**3976**	**0.938**	**4.107**
SVM with LASSO	3600	4375	0.925	4.606
KNN	**3313**	**3861**	**0.942**	**4.317**
KNN with OMP	3520	4084	0.935	4.573
KNN with Boruta	3765	4327	0.927	4.834
KNN with LASSO	3484	4029	0.936	4.544
Adaboost	3214	3991	0.938	4.013
Adaboost with OMP	3074	**3688**	**0.947**	**3.820**
Adaboost with Boruta	3884	4687	0.914	4.949
Adaboost with LASSO	**3067**	3854	0.942	3.831
GBDT	3061	3607	0.949	3.972
GBDT with OMP	3163	3685	0.947	4.151
GBDT with Boruta	3094	3632	0.948	4.043
GBDT with LASSO	**2865**	**3387**	**0.955**	**3.755**

contribution of the examined feature to the forecast values. In a nutshell, a red bar indicates an overall positive contribution and a blue bar indicates an overall negative contribution.

There are similarities and differences between these two best-performing models regarding how to generate their forecasts. First of all, the ESGER is the most influential predictor in the forecasts of the best-performing ML models, whereas EC (industrial) is the predictor with the least predictive power in both models. These findings are consistent with the findings of [70] who claim that increasing electricity production has a positive effect on CO_2 emissions. Most of the directions of feature contributions are in line with each other. However, 3 features out of 10, namely total energy supply, energy consumption, and EC (total), have a different role in affecting the forecasts of these ML models when it comes to making a global interpretation. It should be noted that the second most important variable for the ANN model, GDP, is the second least important for the ANN with the Boruta model. These results are in line with studies [33,41,43] that found positive effects of GDP on CO_2. In addition, it is observed that the impact of total energy supply and energy consumption in the two models cannot be neglected, but their direction of contribution is different from each other in these models. For instance, while total energy supply makes an overall negative contribution to the ANN model, its overall role changes in the ANN model with the Boruta model. These differences are to be expected, since the way the two models make use of variables is different, and there might be some errors associated with the models in approximating the true relationship between variables when modelling them. In this regard, we believe that the common findings identified by models are more reliable for interpretation and it is crucial to check whether the results obtained are compatible with the prevailing belief in the energy field before formulating a policy based on these predictions. In Fig. 3 we can see that the proposed interpretable forecasting framework, ANN with Boruta here, leads to a simpler and easier-to-understand explanation for the problem at hand, by making the differences between variable importance values more pronounced. According to this model, EC (total), population, oil consumption, GDP, and EC (industrial) are the features with low predictive power in forecasting CO_2 emissions in Turkey. These features are the ones with a low correlation value with CO_2 emissions according to the lower part of Fig. 2. However, GDP is found to be the second most important contributor in producing the forecasts of the ANN model.

To further investigate the impact of features on the prediction values,

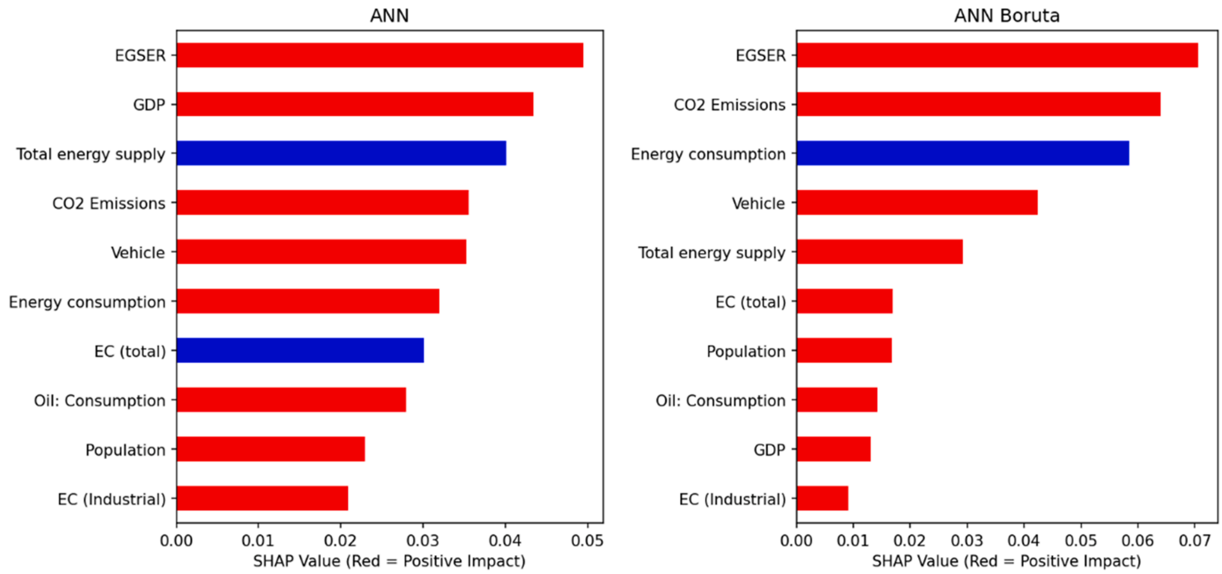

Fig. 3. The average contributions of the features for the CO_2 emissions forecasts of the two best performing models.

a few illustrative SHAP dependence plots are given in Fig. 4. These dependence plots demonstrate the marginal impact of a feature on the outcome of an ML model, and provide an insight into the form of the relationship (if any) between the feature and predicted values, on the one hand, and the feature most likely interacted with, on the other. In Fig. 4, the change in a feature value is represented by the x-axis, and the corresponding Shapley value of that feature is denoted by the y-axis. The colour relates to the other feature that may interact with the feature under investigation. Here we present the dependence plots belonging to only the four most important features detected by the ANN model for CO_2 emissions.

Interpreting the most influential feature, EGSER, found by the two

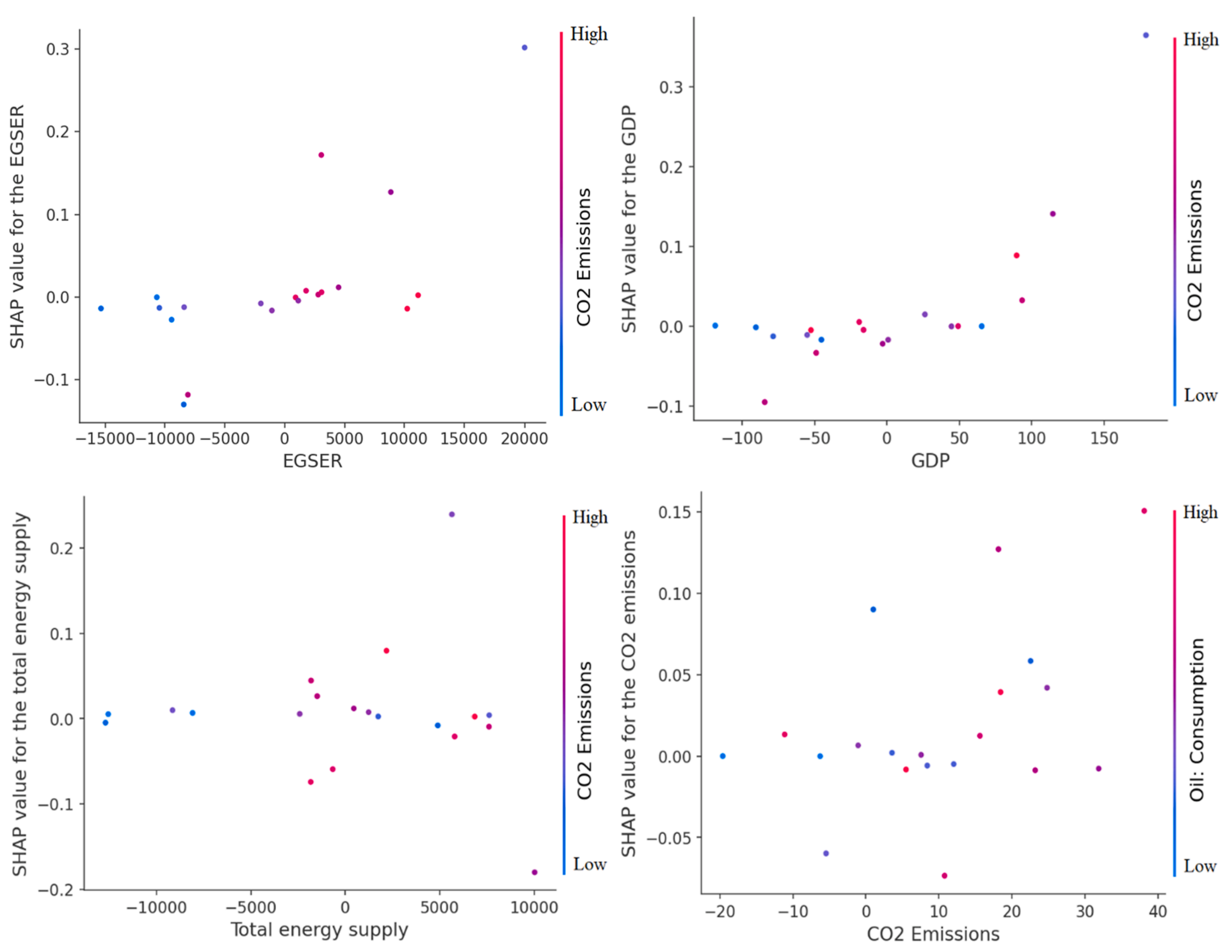

Fig. 4. The dependence plots for CO_2 emissions forecasts.

best-performing models, we can say that there is a trend in the form of a combination of constant and linear trends. Generally speaking, a decrease in EGSER is accompanied by a decrease in CO_2 emissions at time (t-1), and this corresponds to the case where the effect of EGSER on the predictions is very limited. However, the values of EGSER increase with the increases in the value of CO_2 emissions, and this corresponds to the case where the contributions of EGSER to the predictions become more apparent.

The impact of GDP on the predictions is limited in most of the data for the period examined. However, in the years when the change in GDP was positive, it makes some contributions that will increase the prediction values of CO_2 emissions. Even though the average impact of the total energy supply is found to be negative in Fig. 3, unlike the expectations, as can be seen from the bottom left graph in Fig. 4, it is hard to claim a positive or negative relationship between the prediction value and this variable. Depending on the year to be forecasted, the role of this variable may be positive, negative, or neutral. From the SHAP dependence plot at the bottom right of Fig. 4, it can be said that the increase in the amount of CO_2 emissions at time (t-1) may often lead to the expectation of an increase in CO_2 emissions in the following year.

In addition to global explanations, SHAP also offers local explanation tools to demonstrate how much a feature increases or decreases the prediction in a particular instance. For this purpose, we used a waterfall plot presented in Fig. 5 to explain the prediction made for CO_2 emissions in 2020 by the ANN model. The x-axis in the waterfall plot contains the contribution values representing how much each feature moves from the expected value to the final model output. The expected or the base value at the bottom of the graph corresponds to the prediction made by the model if no feature value is known for a particular instance. The y-axis shows the feature values sorted by their contributions and the change in the value of each feature for the examined year. The coloured horizontal bars in the graph denote the amount of the negative (blue) and positive (red) contributions for each feature. The final predicted value is at the top of the graph. Looking at Fig. 5, it can be deduced that all features except the population decrease the prediction value for 2020 by making negative contributions to the expected value. EC (total) and EGSER are the two most important drivers in lowering the prediction value of the model. These findings are in line with our expectations given the outbreak of the COVID-19 pandemic. In 2020, when the social and economic impact of COVID-19 was extremely high, there were high levels of decreases in electricity generation and electricity consumption, especially in parallel with the slowdown and interruptions in industrial production. Fig. 5 shows that the expectation of a decrease in the amount of CO_2 emissions occurring due to these decreases is correctly identified by the model used.

Fig. 6 presents the contribution of each predictor to the prediction made in 2018. Comparing these two figures, we can conclude that the behaviour of the model and the roles of the predictors are different before and after COVID-19. As can be seen from this figure, the expectation produced by the model for the pre-COVID-19 CO_2 emission value is in the direction of increase. The increasing trend from the amount of CO_2 emissions in the previous year continues and makes the biggest contribution to the prediction. The increases observed in GDP, Oil: Consumption, total energy supply, and EC (total) for this year caused our model to predict an increase in CO_2 emissions for the year 2018.

Regarding the interpretation of the forecasts of ML models for the total energy consumption in Turkey, Fig. 7 presents the global bar plot of the two best-performing models, namely GBDT with LASSO, and GBDT. As the figure shows, electricity supply (EGSER) is by far the most influential feature confirmed by both models, whereas the population is the predictor with the least predictive power in both models. Therefore, EGSER reveals itself to be the variable with the highest predictive power among the other variables examined for forecasting CO_2 emissions and energy consumption in Turkey. Considering the close relationship between electricity supply and energy consumption, the fact that ESGER is the variable with the most predictive power in energy consumption by both models and is within the expectations. All of the directions of feature contributions are the same in both models, even though a few differences are observed in order of importance.

Fig. 8 provides more detailed information about the marginal impact of the significant features. The dependence plots belonging to the four most important features detected by GBDT with LASSO are given in Fig. 8. In this figure, we observe that EGSER has a clear positive linear relationship with energy consumption. When we consider the most interacted feature with EGSER, energy consumption at time (t-1), we see that this linear relationship continues to exist in the lagged variable of interest. The dependence plot of oil consumption provides further evidence regarding the negative relationship between this variable and energy consumption. It is well known that the share of oil-based consumption in total energy consumption in Turkey has shown a decreasing trend in the last 30 years. Depending on this, it may be claimed that our model detected this trend to help it in producing its final predictions. Regarding the dependence plot of vehicles, there is a decreasing contribution to the predictions of the energy consumption, especially in the case of decreases in the number of vehicles. This result is in line with the expectation. For the dependence plot of CO_2 emissions, there is no obvious functional relationship to be mentioned. Most data points are clustered around the zero value of SHAP on the y-axis, which means that

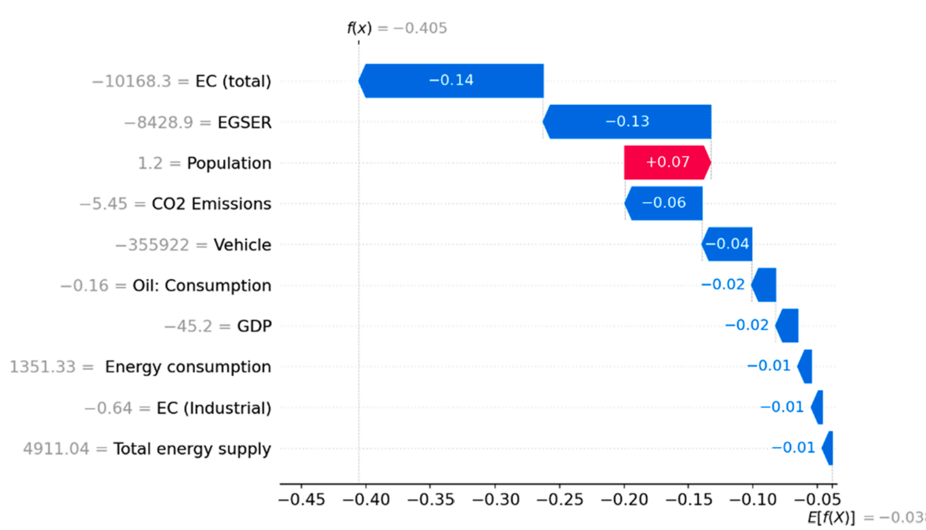

Fig. 5. The local explanation for the forecast of CO_2 emissions in 2020 by the ANN model.

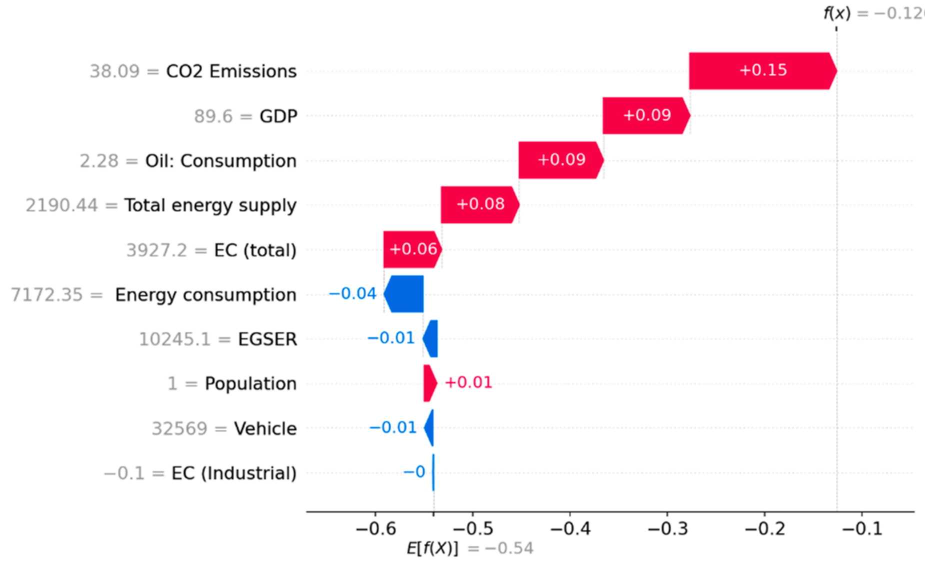

Fig. 6. The local explanation for the forecast of CO_2 emissions in 2018 by the ANN model.

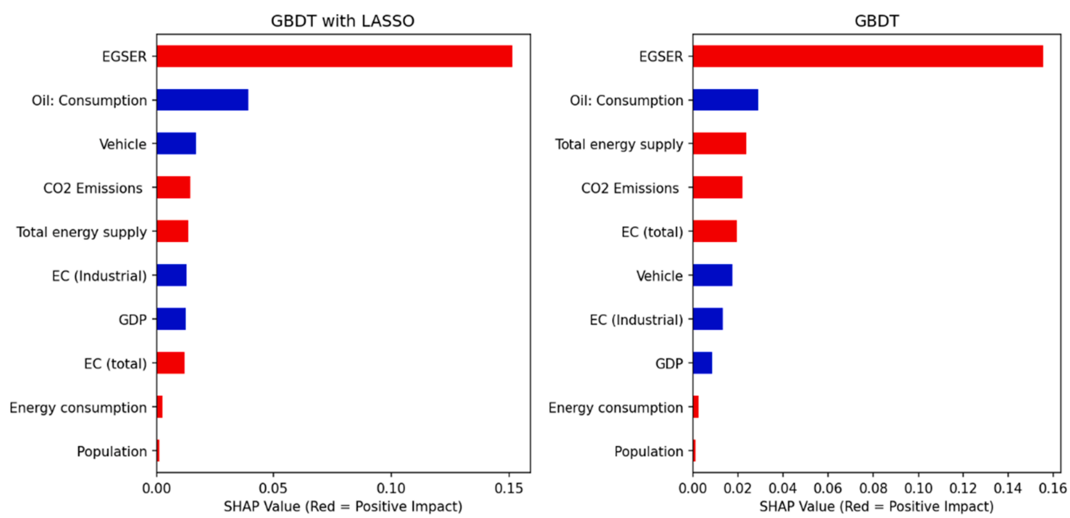

Fig. 7. The average contributions of the features for the energy consumption forecasts of the two best performing models.

the effect of this variable on the prediction is weak.

To pinpoint the role of each predictor locally, we choose again the prediction in 2020 for illustration. Fig. 9 shows the waterfall plot for this year. This figure tells us that only 4 out of 10 predictors have an impact on the prediction of 2020, and EGSER is the most impactful variable. The decrease in the value of EGSER leads to a significant decrease in the prediction. As a result of the decrease in electricity generation due to COVID-19, our model expects a decrease in energy consumption in 2020. Oil consumption and total energy supply have a small positive effect on the prediction. Lastly, the predicted value for this year is lower than the expected value of the model, which means that the model expects a significant reduction in energy consumption for 2020. Fig. 10 provides a local explanation for 2018 by the GBDT with the LASSO model. As expected, the effects of the predictors on the forecasts made before and after COVID-19 have changed dramatically. As expected, electricity generation has increased significantly, making a positive contribution to the prediction of energy consumption for 2018. The values of all predictors except oil consumption have increased, causing a rise in the expectation of energy consumption for this year. The prediction value for 2018 is considerably higher than the expected value of the model.

8. Conclusion

This study constructed a total of 48 forecasting models, each of which is a combination of feature selection techniques and ML models, to determine the best one for energy consumption and CO_2 emissions for Turkey and revealed the relationship between the predictors and predictions. The literature on energy consumption and CO_2 emissions can be divided into two parts. The first part tries to find the relationship between variables by traditional statistical and econometric models without making predictions for the future values of the critical variables. The other part is only concerned with forecasting the variable of interest as accurately as possible. However, in addition to obtaining highly accurate predictions, knowing how a prediction is made and determining whether general patterns exist between outcome variables and covariates is of vital importance for authorised persons and institutions aiming at developing policies effectively when solving problems they have encountered or are likely to encounter. Hence, this study proposed an interpretable forecasting framework that combines feature selection techniques with ML models to assign Shapley values to the variables with high prediction power, to establish the link between two different methodologies used in the field of energy.

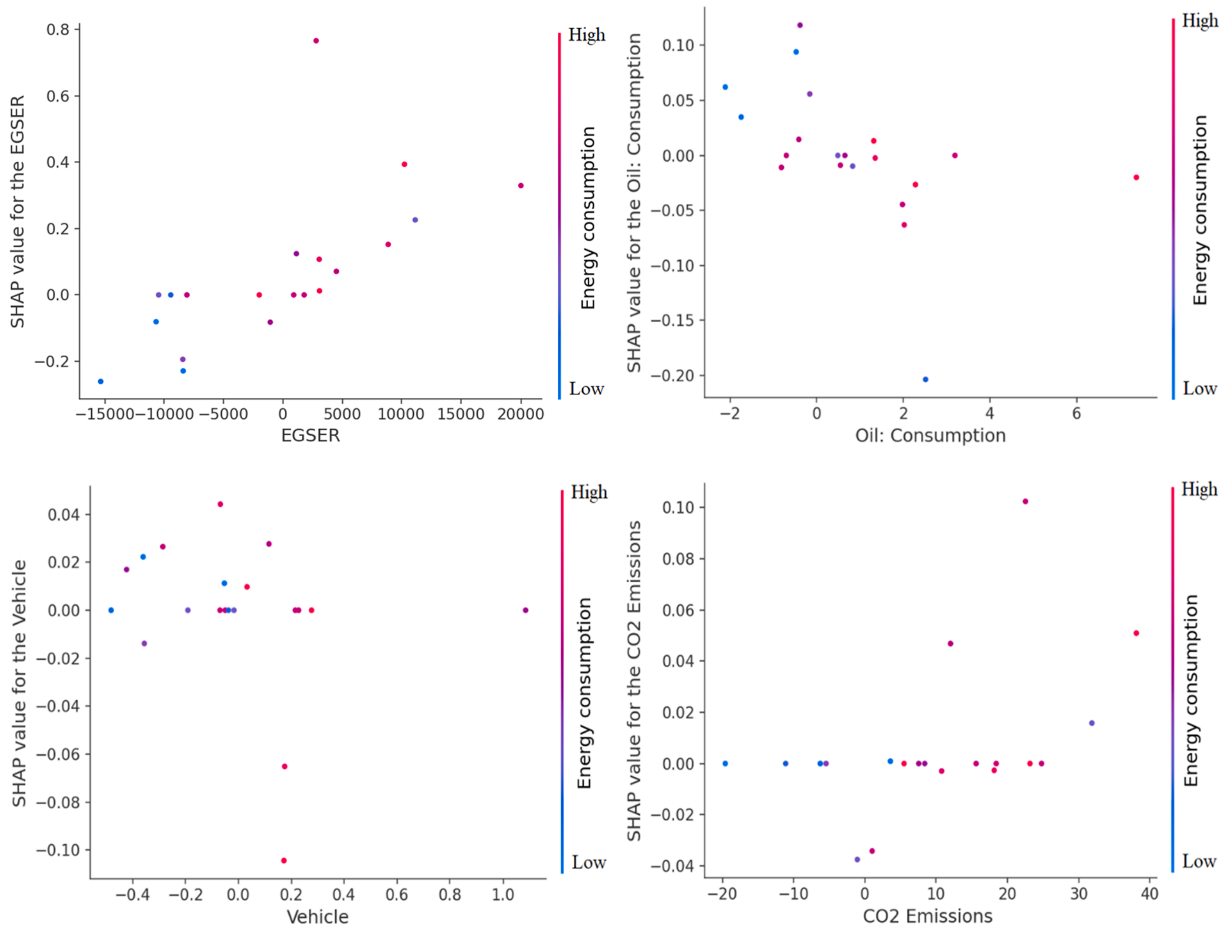

Fig. 8. The dependence plots for energy consumption forecasts.

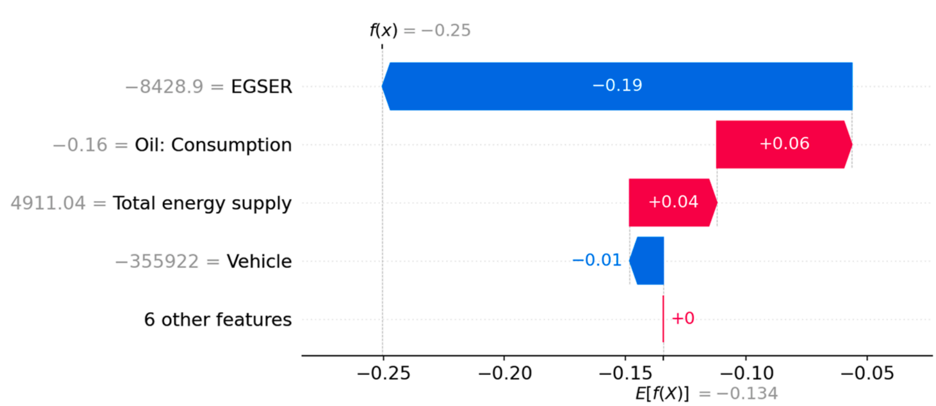

Fig. 9. The local explanation for the forecast of energy consumption in 2020 by GBDT with the LASSO model.

Although the traditional models are transparent about how they work, they may suffer from an inability to provide predictions as good as the ones that ML models offer. These two modelling techniques can be seen as having opposite advantages. However, the methods aiming to open the black-box nature of ML models have recently attracted considerable attention from researchers. Among these methods, SHAP is the most widely used probably due to being an agnostic method and having some valuable theoretical properties. In this study, we have proposed to employ the feature selection techniques coupled with the SHAP method to offer simpler explanations regarding the predictions of energy consumption and CO_2 emissions in Turkey. Based on our empirical data, the following conclusions may be drawn:

- The identified best ML model varies depending on the dataset examined. For CO_2 emissions data, the superiority of ANN models is evident compared to other ML models. However, for energy consumption, GBDT models provide more accurate predictions.
- GDP and EC (industrial) are features with low predictive power in forecasting CO_2 emissions in Turkey.
- EGSER is found to be the variable with the highest predictive power among the other variables examined for forecasting CO_2 emissions and energy consumption in Turkey. This result is plausible considering that Turkey's electricity generation and greenhouse gas emissions have increased by 430 % and 138 %, respectively, in the last 30 years.

model.

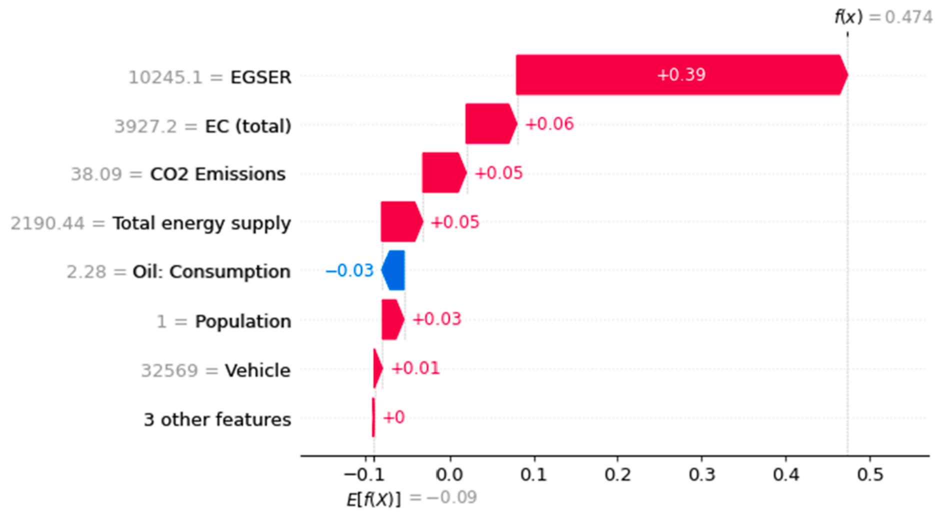

Fig. 10. The local explanation for the forecast of energy consumption in 2018 by GBDT with the LASSO model.

- There is a decreasing contribution to the predictions of energy consumption, especially in cases of decreases in the number of vehicles. This result is in line with the findings of the previous studies [21,71].
- Different from the classical econometric models, certain observations can be explained thanks to the modelling approach utilised. For example, the effect of the COVID-19 pandemic on energy consumption and CO_2 is examined in the year in which it was widespread in the world. It is found that the role of the predictors examined before and after the pandemic has changed considerably.
- In 2020, when COVID-19 first appeared in Turkey, the contribution of all variables on CO_2 except population decreased. EC (total) and EGSER are the two most important drivers in lowering the prediction value of the model. This has shown parallelism with the findings of the previous papers [71–74] examining the effects of COVID-19. This indicates that confidence can be placed in this newly developed methodology.
- Simple and easy-to-use explanations for energy consumption and CO_2 emissions in Turkey can be derived from the methodology utilised in this paper. Since this methodology is general, it can be applied to other variables related to energy or other objects of study to offer simple explanations both locally and globally.

The limitation of the study is the availability of the number of years included in the analysis. Especially when modelling with ML techniques, more observations are needed compared to traditional models to ensure that the results obtained are more consistent and reliable. For future directions, applying interpretable or explainable ML methods to different energy applications can contribute to the expansion of the literature in this area. Also, various statistical tests can be developed to increase the reliability of the results obtained from interpretable ML methods. Thus, besides questioning whether the findings obtained are compatible with the nature of the investigated event, whether they are statistically significant can also be questioned.

CRediT authorship contribution statement

Serkan Aras: Conceptualization, Methodology, Investigation, Validation, Writing – original draft, Writing – review & editing. **M. Hanifi Van:** Conceptualization, Data curation, Investigation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Çodur MY, Ünal A. An estimation of transport energy demand in Turkey via artificial neural networks. Promet - Traffic - Traffico 2019;31:151–61. https://doi.org/10.7307/ptt.v31i2.3041.
[2] Altan İM. Investigation of the environmental impact of growth, energy consumption and financial development in developed and emerging countries: the case of Sweden and Pakistan. Bus Manag Stud An Int J 2021;9:18–31. https://doi.org/10.15295/bmij.v9i1.1633.
[3] Eygü H, Kılınç A. The Analysis of the Effect of Energy Consumption and Carbon-Dioxide Oscillation on Economic Growth: Panel Data Analysis Approach. J Soc Sci Inst 2021;9:268–80. https://doi.org/10.53586/susbid.995759.
[4] Türkoğlu Ü. Çin'de CO2 emisyonu, ekonomik büyüme ve enerji tüketimi arasındaki nedensellik ilişkisi. Çankaya 2021.
[5] Sun W, Liu M. Prediction and analysis of the three major industries and residential consumption CO2 emissions based on least squares support vector machine in China. J Clean Prod 2016;122:144–53. https://doi.org/10.1016/j.jclepro.2016.02.053.
[6] Sechzer PH, Egbert LD, Linde HW, Cooper DY, Dripps RD, Price HL. Effect of carbon dioxide inhalation on arterial pressure, ECG and plasma catecholamines and 17-OH corticosteroids in normal man. J Appl Physiol 1960;15:454–8. https://doi.org/10.1152/jappl.1960.15.3.454.
[7] Dong H, Xue M, Xiao Y, Liu Y. Do carbon emissions impact the health of residents? Considering China's industrialization and urbanization. Sci Total Environ 2021;758:1–14. https://doi.org/10.1016/j.scitotenv.2020.143688.
[8] Pao HT, Fu HC, Tseng CL. Forecasting of CO2 emissions, energy consumption and economic growth in China using an improved grey model. Energy 2012;40:400–9. https://doi.org/10.1016/J.ENERGY.2012.01.037.
[9] World Bank. Gross domestic product 2020. Washington, DC: 2020.
[10] Energy Policy Review. Turkey 2021 Energy Policy Review. 2021.
[11] Turkstat. Greenhouse Gas Emissions Statistics. 2022 2022.
[12] Ribeiro MT, Singh S, Guestrin C. "Why should i trust you?" Explaining the predictions of any classifier. Proc ACM SIGKDD Int Conf Knowl Discov Data Min 2016;13-17-August-2016:1135–44. https://doi.org/10.1145/2939672.2939778.
[13] Carvalho DV, Pereira EM, Cardoso JS. Machine Learning Interpretability: A Survey on Methods and Metrics. Electronics 2019;8(8):832.
[14] Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy 2020;23(1):18.
[15] Lundberg SM, Allen PG, Lee S-I. A Unified Approach to Interpreting Model Predictions. In: Proc. 31st Int. Conf. neural Inf. Process. Syst; 2017. p. 4768–77. https://doi.org/10.5555/3295222.
[16] Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: An overview of interpretability of machine learning. Proc - 2018 IEEE 5th Int Conf

Data Sci Adv Anal DSAA 2018 2019:80–9. https://doi.org/10.1109/DSAA.2018.00018.

[17] Sikder M, Wang C, Yao X, Huai Xu, Wu L, KwameYeboah F, et al. The integrated impact of GDP growth, industrialization, energy use, and urbanization on CO2 emissions in developing countries: Evidence from the panel ARDL approach. Sci Total Environ 2022;837:155795.

[18] Fan Y, Liu LC, Wu G, Wei YM. Analyzing impact factors of CO2 emissions using the STIRPAT model. Environ Impact Assess Rev 2006;26:377–95. https://doi.org/10.1016/j.eiar.2005.11.007.

[19] Liu Y, Hao Y. The dynamic links between CO2 emissions, energy consumption and economic development in the countries along "the Belt and Road". Sci Total Environ 2018;645:674–83. https://doi.org/10.1016/j.scitotenv.2018.07.062.

[20] Jiemin H, Chen W. The impact of private sector energy investment, innovation and energy consumption on China's carbon emissions. Renew Energy 2022;195:1291–9. https://doi.org/10.1016/J.RENENE.2022.06.131.

[21] Ağbulut Ü. Forecasting of transportation-related energy demand and CO2 emissions in Turkey with different machine learning algorithms. Sustain Prod Consum 2022;29:141–57. https://doi.org/10.1016/j.spc.2021.10.001.

[22] Llanos C, Kristjanpoller W, Michell K, Minutolo MC. Causal treatment effects in time series: CO2 emissions and energy consumption effect on GDP. Energy 2022;249:1–20. https://doi.org/10.1016/j.energy.2022.123625.

[23] Kankal M, Uzlu E. Neural network approach with teaching–learning-based optimization for modeling and forecasting long-term electric energy demand in Turkey. Neural Comput Appl 2017;28:737–47. https://doi.org/10.1007/s00521-016-2409-2.

[24] Sakkas N, Yfanti S, Daskalakis C, Barbu E, Domnich M. Interpretable Forecasting of Energy Demand in the Residential Sector. Energies 2021;14:6568. https://doi.org/10.3390/EN14206568.

[25] Wei N, Li C, Peng X, Zeng F, Lu X. Conventional models and artificial intelligence-based models for energy consumption forecasting: A review. J Pet Sci Eng 2019;181:106187. https://doi.org/10.1016/j.petrol.2019.106187.

[26] Mitrentsis G, Lens H. An interpretable probabilistic model for short-term solar power forecasting using natural gradient boosting. Appl Energy 2022;309:118473. https://doi.org/10.1016/J.APENERGY.2021.118473.

[27] Antipov EA, Pokryshevskaya EB. Interpretable machine learning for demand modeling with high-dimensional data using Gradient Boosting Machines and Shapley values. J Revenue Pricing Manag 2020;19:355–64. https://doi.org/10.1057/S41272-020-00236-4/TABLES/5.

[28] Park H, Park DY. Comparative analysis on predictability of natural ventilation rate based on machine learning algorithms. Build Environ 2021;195:107744. https://doi.org/10.1016/J.BUILDENV.2021.107744.

[29] Ben JS, Mefteh-Wali S, Viviani JL. Forecasting gold price with the XGBoost algorithm and SHAP interaction values. Ann. Oper Res 2021:1–21. https://doi.org/10.1007/S10479-021-04187-W/FIGURES/4.

[30] Pokharel S, Sah P, Ganta D. Improved Prediction of Total Energy Consumption and Feature Analysis in Electric Vehicles Using Machine Learning and Shapley Additive Explanations Method. World Electr Veh J 2021;12:94. https://doi.org/10.3390/WEVJ12030094.

[31] Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. Npj Digit Med 2021;4:1–5. https://doi.org/10.1038/s41746-020-00372-6.

[32] Chen S-X, Wang X-K, Zhang H-y, Wang J-Q, Peng J-J. Customer purchase forecasting for online tourism: A data-driven method with multiplex behavior data. Tour Manag 2021;87:104357.

[33] Ang JB. Economic development, pollutant emissions and energy consumption in Malaysia. J Policy Model 2008;30:271–8. https://doi.org/10.1016/j.jpolmod.2007.04.010.

[34] Zhang XP, Cheng XM. Energy consumption, carbon emissions, and economic growth in China. Ecol Econ 2009;68:2706–12. https://doi.org/10.1016/j.ecolecon.2009.05.011.

[35] Ajmi AN, Hammoudeh S, Nguyen DK, Sato JR. On the relationships between CO2 emissions, energy consumption and income: The importance of time variation. Energy Econ 2013;49:629–38. https://doi.org/10.1016/j.eneco.2015.02.007.

[36] Grossman GM, Krueger AB. Environmental Impacts of a North American Free Trade Agreement. National Bureau of Economic Research. NBER Work Aper Ser 1991. https://doi.org/10.3386/W3914.

[37] Grossman GM, Krueger AB. Economic growth and the environment. Q J Econ 1995;110:353–77. https://doi.org/10.2307/2118443.

[38] Van MH, Sadradin HF. The Effect of Technological Innovation on Co2 Emissions in OECD Countries: Using Panel Regression. Int J Contemp Econ Adm Sci 2021;2:438–53. https://doi.org/10.5281/zenodo.5831707.

[39] Arvin MB, Pradhan RP, Norman NR. Transportation intensity, urbanization, economic growth, and CO2 emissions in the G-20 countries. Util Policy 2015;35:50–66. https://doi.org/10.1016/j.jup.2015.07.003.

[40] Murat YS, Ceylan H. Use of artificial neural networks for transport energy demand modeling. Energy Policy 2006;34:3165–72. https://doi.org/10.1016/j.enpol.2005.02.010.

[41] Pata UK. The effect of urbanization and industrialization on carbon emissions in Turkey: evidence from ARDL bounds testing procedure. Environ Sci Pollut Res 2018;25:7740–7. https://doi.org/10.1007/s11356-017-1088-6.

[42] Le TH. Connectedness between nonrenewable and renewable energy consumption, economic growth and CO2 emission in Vietnam: New evidence from a wavelet analysis. Renew Energy 2022;195:442–54. https://doi.org/10.1016/j.renene.2022.05.083.

[43] Cherni A, Essaber JS. An ARDL approach to the CO2 emissions, renewable energy and economic growth nexus: Tunisian evidence. Int J Hydrogen Energy 2017;42:29056–66. https://doi.org/10.1016/j.ijhydene.2017.08.072.

[44] Narayan PK, Smyth R, Prasad A. Electricity consumption in G7 countries: A panel cointegration analysis of residential demand elasticities. Energy Policy 2007;35:4485–94. https://doi.org/10.1016/j.enpol.2007.03.018.

[45] Limanond T, Jomnonkwao S, Srikaew A. Projection of future transport energy demand of Thailand. Energy Policy 2011;39:2754–63. https://doi.org/10.1016/j.enpol.2011.02.045.

[46] Ofosu-Adarkwa J, Xie N, Javed SA. Forecasting CO2 emissions of China's cement industry using a hybrid Verhulst-GM(1, N) model and emissions' technical conversion. Renew Sustain Energy Rev 2020;130:109945. https://doi.org/10.1016/j.rser.2020.109945.

[47] Mirzaei M, Bekri M. Energy consumption and CO2 emissions in Iran, 2025. Environ Res 2017;154:345–51. https://doi.org/10.1016/j.envres.2017.01.023.

[48] Li H, Lin J, Lei X, Wei T. Compressive strength prediction of basalt fiber reinforced concrete via random forest algorithm. Mater Today Commun 2022;30:1–9. https://doi.org/10.1016/j.mtcomm.2021.103117.

[49] Géron A. Hands-on Machine Learning Scikit-Learn and TensorFlow Concepts. Tools, and Techniques to Build Intelligent Systems 2017.

[50] Zhelavskaya IS, Shprits YY, Spasojevic M. Reconstruction of plasma electron density from satellite measurements via artificial neural networks. Mach Learn Tech Sp Weather, Elsevier 2018:301–27. https://doi.org/10.1016/B978-0-12-811788-0.00012-3.

[51] Zhang N, Yang B, Liu K, Li H, Chen G, Qiu X, et al. Machine Learning in Screening High Performance Electrocatalysts for CO2 Reduction. Small Methods 2021;5(11):2100987.

[52] González PA, Zamarreño JM. Prediction of hourly energy consumption in buildings based on a feedback artificial neural network. Energy Build 2005;37:595–601. https://doi.org/10.1016/J.ENBUILD.2004.09.006.

[53] Lantz B. Machine Learning with R: Expert techniques for predictive modeling. 3rd Edition. BIRMINGHAM - MUMBAI: Packt Publishing Ltd; 2019.

[54] Wen L, Cao Y. Influencing factors analysis and forecasting of residential energy-related CO2 emissions utilizing optimized support vector machine. J Clean Prod 2020;250:1–13. https://doi.org/10.1016/j.jclepro.2019.119492.

[55] Dong B, Cao C, Lee SE. Applying support vector machines to predict building energy consumption in tropical region. Energy Build 2005;37:545–53. https://doi.org/10.1016/j.enbuild.2004.09.009.

[56] Freund Y, Schapire RE. Experiments with a new boosting algorithm. Proc. 13th Int. Conf. om Mach. Learn., Bari, Italy: 1996, p. 148–56.

[57] Drucker H. Improving Regressors using Boosting Techniques. In: Kaufmann M, editor. ICML Proc. Lille: Fourteenth Int. Conf. Mach. Learn; 1997. p. 107–15.

[58] Friedman JH. Greedy Function Approximation: A Gradient Boosting Machine on JSTOR. Ann Stat 2001;29:1189–232.

[59] Shapley LS. A Value for n-Person Games. Contrib. to Theory Games (AM-28), Vol. II, Princeton University Press; 1953, p. 307–18. https://doi.org/10.1515/9781400881970-018.

[60] Štrumbelj E, Kononenko I. An Efficient Explanation of Individual Classifications using Game Theory. J Mach Learn Res 2010;11:1–18.

[61] Tibshirani R. Regression Shrinkage and Selection Via the Lasso. J R Stat Soc Ser B 1996;58:267–88. https://doi.org/10.1111/J.2517-6161.1996.TB02080.X.

[62] Pati YC, Rezaiifar R, Krishnaprasad PS. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: Proc. 27th Asilomar Conf. signals, Syst. Comput., Publ by IEEE; 1993. p. 40–4. https://doi.org/10.1109/ACSSC.1993.342465.

[63] Mallat SG, Zhang Z. Matching Pursuits With Time-Frequency Dictionaries. IEEE Trans Signal Process 1993;41:3397–415. https://doi.org/10.1109/78.258082.

[64] Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. J Stat Softw 2010;36:1–13. https://doi.org/10.18637/JSS.V036.I11.

[65] IEA. World Energy Outlook 2021 – Analysis - IEA 2022. https://www.iea.org/reports/world-energy-outlook-2021 (accessed May 6, 2022).

[66] Turkstat. Number of Road Motor Vehicles. 2021 2021. https://data.tuik.gov.tr/Bulten/Index?p=Road-Motor-Vehicles-December-2020-37410 (accessed May 6, 2022).

[67] World Bank. World Development Indicators | DataBank 2022. https://databank.worldbank.org/source/world-development-indicators (accessed May 6, 2022).

[68] Sundararajan M, Najmi A. The Many Shapley Values for Model Explanation. In: Proc. 37th Int. Conf. Mach. Learn; 2020. p. 9269–78.

[69] Hyndman RJ. Another look at forecast-accuracy metrics for intermittent demand. Foresight Int J Appl Forecast 2006;4:43–6.

[70] Ike GN, Usman O, Sarkodie SA. Testing the role of oil production in the environmental Kuznets curve of oil producing countries: New insights from Method of Moments Quantile Regression. Sci Total Environ 2020;711:135208. https://doi.org/10.1016/j.scitotenv.2019.135208.

[71] Bazzo Vieira JP, Vieira Braga CK, Pereira RHM. The impact of COVID-19 on air passenger demand and CO2 emissions in Brazil. Energy Policy 2022;164:112906. https://doi.org/10.1016/j.enpol.2022.112906.

[72] Sözen A, İzgeç MM, Kırbaş İ, Kazancıoğlu FŞ, Tuncer AD. Overview, modeling and forecasting the effects of COVID-19 pandemic on energy market and electricity

demand: a case study on Turkey. Energy Sources, Part A Recover Util Environ Eff 2021:1–16. https://doi.org/10.1080/15567036.2021.1910756.
[73] Andreoni V. Estimating the European CO2 emissions change due to COVID-19 restrictions. Sci Total Environ 2021;769:145115. https://doi.org/10.1016/j.scitotenv.2021.145115.
[74] Smith LV, Tarui N, Yamagata T. Assessing the impact of COVID-19 on global fossil fuel consumption and CO2 emissions. Energy Econ 2021;97:105170. https://doi.org/10.1016/j.eneco.2021.105170.