Journal of Environmental Management 359 (2024) 121040

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

A comprehensive analysis of key factors' impact on environmental performance: Evidence from Globe by novel super learner algorithm

Mustafa Tevfik Kartal [a,b,c,d,e,*], Özer Depren [d,f], Serpil Kılıç Depren [g]

[a] *Department of Economics and Management, Khazar University, Baku, Azerbaijan*
[b] *Department of Finance and Banking, European University of Lefke, Lefke, Northern Cyprus, TR-10 Mersin, Türkiye*
[c] *Adnan Kassar School of Business, Lebanese American University, Beirut, Lebanon*
[d] *Clinic of Economics, Azerbaijan State University of Economics (UNEC), Baku, Azerbaijan*
[e] *GUST Center for Sustainable Development, Gulf University for Science and Technology, Kuwait*
[f] *Customer Experience Research Lab, Yapı Kredi Bank, İstanbul, Türkiye*
[g] *Department of Statistics, Yildiz Technical University, İstanbul, Türkiye*

A R T I C L E I N F O

JEL classification:
C45
Q48
Q56

Keywords:
Environmental performance
Machine learning
Super learner
Global sample

A B S T R A C T

This study aims to analyze comprehensively the impact of different economic and demographic factors, which affect economic development, on environmental performance. In this context, the study considers the Environmental Performance Index as the response variable, uses GDP per capita, tariff rate, tax burden, government expenditure, inflation, unemployment, population, income tax rate, public debt, FDI inflow, and corporate tax rate as the explanatory variables, examines 181 countries, performs a novel Super Learner (SL) algorithm, which includes a total of six machine learning (ML) algorithms, and uses data for the years 2018, 2020, and 2022. The results demonstrate that (i) the SL algorithm has a superior capacity with regard to other ML algorithms; (ii) gross domestic product per capita is the most crucial factor in the environmental performance followed by tariff rates, tax burden, government expenditure, and inflation, in order; (iii) among all, the corporate tax rate has the lowest importance on the environmental performance followed by also foreign direct investment, public debt, income tax rate, population, and unemployment; (iv) there are some critical thresholds, which imply that the impact of the factors on the environmental performance change according to these barriers. Overall, the study reveals the nonlinear impact of the variables on environmental performance as well as their relative importance and critical threshold. Thus, the study provides policymakers valuable insights in re-formulating their environmental policies to increase environmental performance. Accordingly, various policy options are discussed.

1. Introduction

In an era dominated by the urgent need to address environmental challenges, evaluating a country's environmental performance has become a critical indicator of its dedication to sustainable practices. The complex dynamics of climate change, resource exhaustion, and biodiversity loss have underscored the necessity of understanding the diverse factors shaping a country's ecological stance. Thus, understanding and actively managing the factors influencing a country's environmental performance has become an important point for countries and policymakers (Kılıç Depren et al., 2022; Kartal et al., 2023a).

The significance of a country's environmental performance extends beyond being only a measurement; it stands for a statement of a country's commitment to sustainable development and the preservation of the global ecosystem. In other words, the environmental performance of a country is a crucial indicator of its responsibility toward mitigating climate change, conserving natural resources, and safeguarding biodiversity (Hassan et al., 2023). A robust environmental performance reflects a country's dedication to reducing its ecological footprint and fostering practices that promote long-term ecological resilience. Although a positive environmental performance is indicative of a country's commitment to international agreements and frameworks aimed at addressing global environmental issues, the geographic conditions may shape a country's environmental baseline. Thus, geographical factors (climate, topography, natural resources, etc.) have been extensively important for a country. More specifically, it is revealed that ecosystem productivity and biodiversity are linked to geographic conditions, and they influence a country's environmental resilience and sustainability (Sala et al., 2000). However, it's important to note that environmental performance goes beyond geographic

* Corresponding author. Department of Economics and Management, Khazar University, Baku, Azerbaijan.

https://doi.org/10.1016/j.jenvman.2024.121040
Received 19 February 2024; Received in revised form 13 April 2024; Accepted 27 April 2024

determinants.

According to the study of Pata et al. (2023a), it is emphasized the role of governance, policy effectiveness, and international cooperation in driving positive environmental outcomes. Countries with strong institutional frameworks and proactive policies tend to outperform others in environmental sustainability, irrespective of their natural endowment. Furthermore, empirical studies have demonstrated the impact of international agreements on environmental behavior. It is also emphasized that countries participating in global environmental agreements exhibit greater compliance and environmental performance, reflecting their commitment to international cooperation (Rehman et al., 2023). Thus, it can be said that countries that actively work towards achieving and surpassing environmental targets demonstrate their role as responsible global citizens, contributing to a collective effort to address climate change and protect the planet's delicate ecological balance.

From a socio-economic perspective, a high level of environmental performance has tangible benefits for a country. It fosters a healthier and more sustainable living environment for its citizens, leading to improved public health and reduced healthcare costs associated with environmental pollution (Shang et al., 2022; Yang et al., 2022). Furthermore, a commitment to sustainable practices attracts businesses and investments as environmentally conscious industries and consumers increasingly prioritize countries with a demonstrated commitment to environmental management. The interplay between fiscal policies and environmental results is a nuanced subject warranting exploration. Recent studies indicate that countries with lower tariff rates on green technologies and a more progressive tax burden tend to exhibit a more favorable environment for sustainable development (Hao et al., 2021; Sharif et al., 2023). Also, the allocation of government expenditure plays a pivotal role in determining a country's commitment to environmental conservation and sustainable development (Zhang et al., 2017; Koçak and Ulucak, 2019). In addition, population dynamics, GDP per capita, unemployment, and inflation collectively contribute to the socio-economic context influencing a country's environmental stance. Countries, that have a steadily increasing GDP per capita and lower unemployment levels, tend to exhibit more robust environmental performance metrics (Ullah et al., 2020; Ulussever et al., 2023). Finally, FDI inflow represents a double-edged sword, capable of either bolstering or compromising a country's environmental objectives (Ullah et al., 2023). Notably, countries receiving higher FDI inflows tend to experience an increase in industrial emissions, but this impact can be mitigated by stringent environmental regulations (Essandoh et al., 2020).

Governments are under more pressure to do better in terms of environmental performance. According to the 2030 SDGs Agenda, countries have a responsibility to educate their citizens about their environmental policies for reducing pollution and managing natural resources, thereby safeguarding the sustainability of their own countries (United Nations, 2015). Therefore, environmental measurements are necessary for governments to demonstrate the effectiveness of their policies and programs, assess progress in terms of noticeable impacts and improvements to environmental health, and ensure the relevance of policy planning and decision-making processes.

To comprehensively evaluate the environmental performance of the countries, Yale and Columbia Universities propose the EPI, which is primarily based on a set of measures that are used to assess environmental issues, such as pollution, natural resource management, environmental health, ecosystem quality, and climate change (Wolf et al., 2022). These indicators may include, among other things, water and air quality, waste management, biodiversity, and forest conservation. The 2022 EPI Framework comprises 40 environmental performance indicators distributed across 11 issue categories, which are further aggregated into three overarching policy objectives. The weights assigned to each category represent the percentage contribution to the overall EPI score (Wolf et al., 2022). So, it is possible to compare the environmental performance of various countries by using the EPI because it allows them to evaluate countries' environmental policies, identify areas of weakness, and facilitate reforms (Fu et al., 2020; Khan et al., 2020). In addition to environmental performance, the EPI might consider social and economic issues. This provides a thorough examination of environmental sustainability since environmental performance can be influenced by many factors (e.g., social inequality, poverty, and economic development). So, their inclusion in the assessment is critical.

Considering the emergence of the EPI, the study aims to investigate the impact of the various economic and demographic factors on the environmental performance of the countries. In doing so, the study considers the nonlinear impacts and potential thresholds. Also, the study considers a large scope, including 181 countries, and applies a novel SL algorithm. So, the study obtains a higher estimation capacity, determines the relative importance of the effective factors on environmental performance, and reveals the critical thresholds. In this way, the study provides some contributions to the literature; (i) the study first investigates the relationship between demographics and basic economic indicators versus environmental performance; (ii) instead of using CO_2 emissions, ecological footprint, $PM_{2.5}$, or load capacity factor as an indicator of the environment, the EPI, which is a much more comprehensive indicator consisting of 40 different environmental performance indicators, is used; (iii) the SL approach, which is a novel ensemble ML model, is applied to improve the robustness of the fundamental ML approaches; (iv) rather than examining any single country, the study analyzes a total of 181 countries by using data for the years 2018, 2020, and 2022.

Section 2 constructs the theoretical background and reviews the literature. Section 3 presents the dataset and overview of the methodology. The findings obtained are presented in Section 4. Finally, Section 5 presents the conclusion, policy implications, and further research.

2. Theoretical framework and literature review

Within the existing body of literature, several studies have endeavored to investigate the interconnection between environmental performance and explanatory indicators covering macro/microeconomic dynamics or energy consumption. These analytical studies typically involve economic indicators, political uncertainties, and the production/consumption metrics of different energy sources. Notably, extant literature conventionally employs CO_2 emissions as the primary indicator for measuring environmental performance (e.g., Saidi and Omri, 2020; Bekun, 2024). Also, later studies have used either ecological footprint (e.g., Kartal and Pata, 2023; Dam et al., 2024) or load capacity factor (e.g., Kartal et al., 2023b; Lin and Ullah, 2024). Different from such studies, this study adopts a distinctive approach by employing the EPI as a more comprehensive metric for the environment. Consequently, the analytical framework integrates a holistic index that allows for a nuanced examination of environmental performance across its multifaceted dimensions.

2.1. Theoretical framework

The EKC hypotheses serve as the primary foundation for this study's theoretical framework (Grossman and Krueger, 1991). The EKC is a theoretical framework that explores the relationship between environmental degradation and economic development. The curve suggests that, initially, as a country experiences economic growth, environmental quality deteriorates. However, beyond a certain income threshold, environmental degradation begins to decrease as economic development progresses. Hence, it becomes evident that a robust correlation exists between economic growth and environmental performance.

Various economic indicators, including those related to fiscal policy, taxation, government spending, population dynamics, and economic performance, play crucial roles in shaping the trajectory of the EKC (Hashmi and Alam, 2019; Yuelan et al., 2022; Pata et al., 2023b; Adekoya et al., 2024). In light of these studies emphasizing the EKC theory, variables about fiscal policy, taxation, government spending, population dynamics, and economic performance are also included as explanatory factors in analyzing the EPI.

2.2. Literature review

2.2.1. Economic growth, public debt, and environment

Although focusing solely on economic growth can lead to serious consequences, countries have often prioritized economic efficiency over the needs of the environment (Khan et al., 2021a; Shittu et al., 2021). As countries become wealthier, the influence of individualism becomes more intense than the GDP per capita impact. Improvements in the GDP per capita lead to better environmental performance (Kumar et al., 2019). Chowdhury and Islam (2017) investigate the relationship between EPI and the GDP growth in BRICS countries and the results show a negative relationship between EPI and GDP growth. Pimonenko et al. (2018) conduct a study, that finds a strong correlation between GDP and EPI, indicating that a policy aimed at eco-friendly economic growth can help to improve both economic and environmental performance. Li et al. (2021) provide evidence of the factors that influence the EPI in Asian countries and define an insignificant positive impact of GDP on the EPI. Also, there are up-to-date studies (e.g., Dedeoğlu et al., 2021; Khan et al., 2021b) that examine economic growth and environmental performance relationship using different approaches and the impact of economic growth on the environment differs from country to country.

In addition, a few studies have explored the relationship between public debt and the environment. This is because the relationship between public debt and environmental performance is complex and difficult to determine. Various factors, such as economic growth, resource allocation, and governance mechanisms influence this relationship. Environmental regulatory programs are often hindered due to budget deficits and the growth of public debt (Fodha and Seegmuller, 2014). It is observed that countries with better environmental performance are associated with public debt (Clootens, 2017; Ulman et al., 2021).

2.2.2. Foreign direct investment and environment

The impact of FDI inflows on environmental performance has been a subject of ongoing debate. The lack of consensus on this matter can be attributed to the complex nature of the relationship between FDI and the environment (Hao et al., 2020; Xie et al., 2023). FDI has been found to have a positive and significant impact on environmental performance in developed countries. However, in developing countries, the impact of FDI on environmental performance may be either insignificant or negative. Furthermore, it is imperative to note that the impact of FDI on environmental performance in developed countries manifests heterogeneity when analyzed across various quantiles of environmental performance (Li et al., 2019).

2.2.3. Inflation, unemployment, and environment

Although studies on the relationship between inflation and the environment are limited, some findings about it are available. Inflation is generally considered an economic indicator and may have indirect impacts on the environment. Specifically, high inflation can reduce consumer purchasing power, decrease demand for goods, and hurt future investment expenditures. Also, an increase in prices can diminish the purchasing power of government expenditures. In such a case, the government may increase its spending, resulting in a decrease in demand for goods. Consequently, all these factors reduce aggregate demand and generally improve environmental performance (Ahmad et al., 2021). These findings contradict the study of Chambers (2011), who claims that macroeconomic instability has a positive impact on environmental pollution.

Unemployment is also an essential determinant that affects people's health and the environment (Xin et al., 2023). Kashem and Rahman (2020) make an effort to introduce a new concept, the EPC, to explain the relationship between unemployment and the environment and define a negative correlation between pollution and unemployment in 30 industrialized countries. Scholars posit that the adoption of viable technological solutions holds the potential to mitigate pollution while concurrently preserving or enhancing employment rates within the economy. In the studies by Anser et al. (2021) and Ng et al. (2022), the authors provide evidence for the negative relationship between unemployment and the environment under the EPC scope.

2.2.4. Population and environment

The rapid growth of the world population is often cited as a major driver of environmental degradation. Studies conducted by Kumar et al. (2019), Raza et al. (2021), Le and Hoang (2022) and Stoian et al. (2022) have found that an increase in population causes a degradation impact on the environment. However, population size alone is not the sole determinant of environmental performance. Other non-economic factors, such as lifestyle, consumption patterns, technological advancement, and environmental policies also play important roles. Fu et al. (2020) argue that population growth can promote environmental performance by improving living standards, increasing educational levels, and raising environmental awareness.

3. Methods

3.1. Data

The study aims to model the EPI by considering a set of explanatory variables. In this context, data is collected from two fundamental data sources, which are Heritage (2023) and Yale Center for Environmental Law and Policy (2023) for explanatory variables and EPI, in order. Since the EPI is measured bi-annually, the aggregated dataset is prepared for 2018, 2020, and 2022. For each year, there are 181 countries' data. So, the details of the variables are given in Table 1.

There are 533 observations in the data set, but there are missing data on some explanatory variables due to the inability to collect data. More specifically, ~4% of the total data is missing, which means only 24 observations. For this reason, the percentage of missing data is less than 5%, listwise deletion is applied following the literature (Mirzaei et al., 2022; Liu et al., 2024), and a total of 509 data are included in the analysis. According to the studies in the literature, it can be reached accurate outcomes with this number of observations (Ramezan et al., 2021; Rajput et al., 2023). To mitigate the potential issue of overfitting in the modeling, the dataset is divided into two sub-sets (i) the training set, which consists of 70% by encompassing 373 data, and (ii) the testing set, which comprises the remaining 30% by encompassing 136 data. Furthermore, a 10-fold cross-validation procedure is implemented to

Table 1
Variables.

Symbol	Definition	Unit	Source
EPI	Environmental Performance Index[a]	Index (0–100)	Yale Center for Environmental Law & Policy (2023)
TR	Tariff Rate	%	Heritage (2023)
ITR	Income Tax Rate	%	
CTR	Corporate Tax Rate	%	
TB	Tax Burden	% of GDP	
GE	Government Expenditure	% of GDP	
POP	Population	Millions	
GDPPP	GDP per Capita	USD	
UN	Unemployment	%	
INF	Inflation	%	
FDI	FDI Inflow	Million USD	
PUBD	Public Debt	% of GDP	

Notes.
[a] Denotes the response variable.

ensure the robustness and reliability of the results (Kılıç Depren and Depren, 2021).

3.2. EPI framework

The EPI is a methodology covering 40 environmental indicators within 11 dimensions to provide a data-driven single index value (ranging from 0 to 100) presenting the sustainability level of a country (Wolf et al., 2022). These 11 dimensions and their weights are Climate Change Mitigation (38%), Biodiversity & Habitat (18%), Air Quality (11%), Ecosystem Services (8%), Fisheries (5%), Sanitation & Drinking Water (5%), Acid Rain (4%), Agriculture (4%), Water Resources (3%), Heavy Metals (2%), and Waste Management (2%). In addition, the EPI value is calculated for approximately 180 countries each year by using the environmental indicators in the aforementioned dimensions, which are presented in Appendix 1 in detail.

Although the primary focus of the EPI is on environmental indicators (such as CO_2 Growth Rate, GHG Emissions per Capita, $PM_{2.5}$ Exposure, Unsafe Drinking Water, Wastewater Treatment, etc.), it considers the interdependence of social, economic, and environmental aspects in determining sustainability. In addition, there are no factors about the social and economic issues in the calculation of the EPI directly. The presence of economic and social issues in the EPI reflects a recognition that larger societal and economic factors frequently influence environmental results. For example, factors such as income inequality, access to education and healthcare, and employment opportunities can impact a population's environmental awareness, behavior, and ability to adopt sustainable practices. The EPI aims to provide a comprehensive evaluation of a country's environmental performance, taking into consideration that environmental sustainability is linked to social and economic well-being. Therefore, it doesn't ignore social and economic concerns. Instead, it recognizes sustainability's complexity and the importance of integrated strategies to effectively tackle environmental problems.

The ranking of the EPI in 2022 for 180 countries is visualized in Fig. 1.

In the EPI ranking, countries are colored from green to red indicating the best performer to the worst performer countries regarding environmental performance, respectively. Based on the EPI ranking in 2022, Denmark, United Kingdom, and Finland are in the top 3 places with values of 77.9, 77.7, and 76.5, respectively. Besides, India, Myanmar, and Vietnam are in the bottom 3 places with values of 18.9, 19.4, and 20.1, respectively. In this methodology, higher values of EPI represent better environmental performance while lower values of EPI represent worse environmental performance.

3.3. Empirical methodology

The empirical methodology of the study consists of six steps as displayed in Fig. 2.

In the first step, the basic distributional characteristics of the data, such as mean, median, quartiles, and standard deviations, are given in detail. Six different ML algorithms are employed in the second step. These algorithms are CART, RF, XGB, GBM, SVM, and k-NN. Performing the SL algorithm (Van der Laan et al., 2007; Van der Laan and Rose, 2011), which is an ensemble model combining more than one different ML algorithm, is the third step of the methodology. The SL algorithm follows a structured procedure consisting of seven key steps. Initially, the dataset is divided into training and testing sets to facilitate model evaluation. Subsequently, six distinct base learners are trained on the training data. The test set predictions from each base learner are saved for further analysis. To assess model performance, these predictions are evaluated using the test dataset. Next, a meta-model is employed to derive optimal weights based on the test set predictions. Following this, the base learners are re-trained on the entire training dataset, and their predictions are saved. Finally, the SL combines these predictions using the calculated weights to generate the final ensemble prediction. This systematic approach ensures that the SL leverages the strengths of diverse base learners and effectively integrates their predictions to enhance overall performance. Comparing the model performance metrics and visualizing the actual vs predicted observations of the best model are the fourth and fifth steps of the methodology. The sixth step explores the determination of the most important to less important variables.

3.4. Model performance criteria

The MAE, the RMSE, and the R^2 are used to assess the performance of ML algorithms as formulated in Eqs. (1)–(3):

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

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

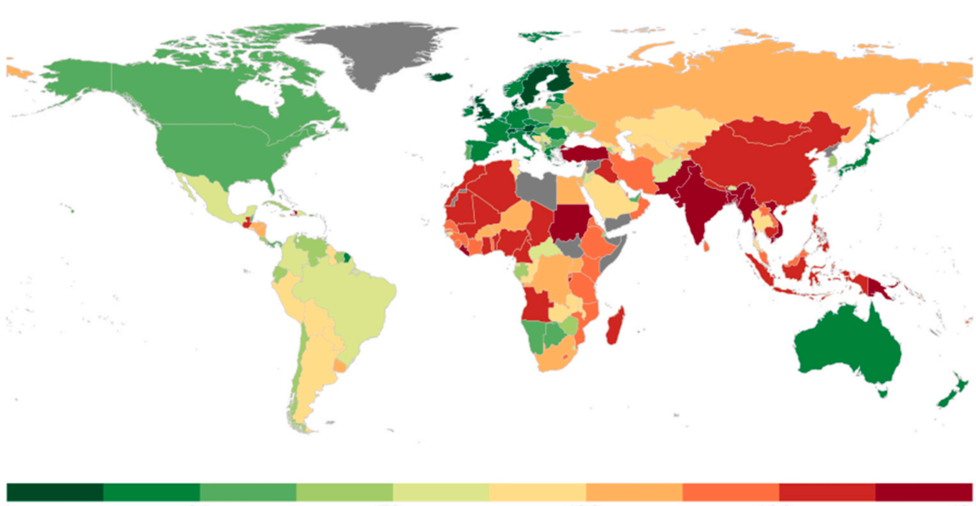


Fig. 1. EPI 2022 Scores of countries.
Source: EPI 2022 Technical Report (Wolf et al., 2022)

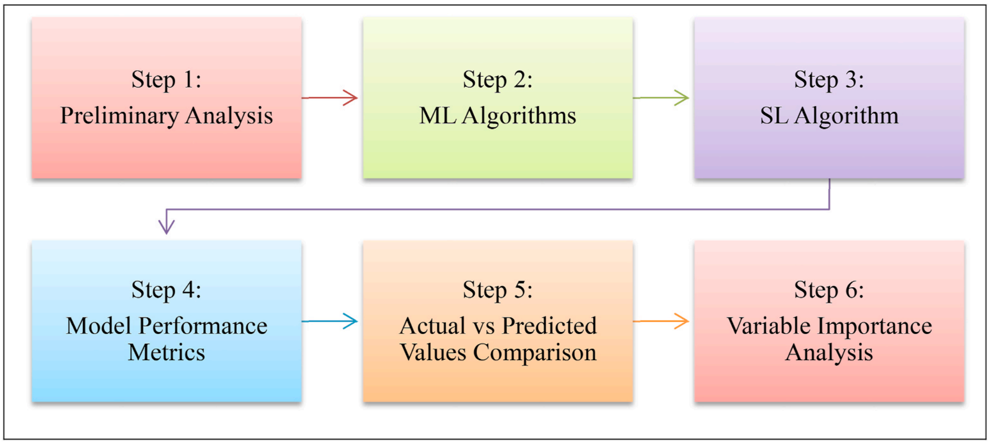


Fig. 2. The steps of empirical methodology.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y}_i)^2} \tag{3}$$

In the context of a test set with n samples, where y_i represents the actual values and $\hat{y}_i$ represents the predicted values, lower RMSE, and MAE values suggest reduced error levels, indicating greater accuracy (Willmott and Matsuura, 2005). The fitting effectiveness improves as the R^2 value approaches 1.

4. Results

4.1. Preliminary Statistics

The basic distributional characteristics are visualized in Fig. 3. Also, the central tendency and variation statistics and tests for normal distribution are presented.

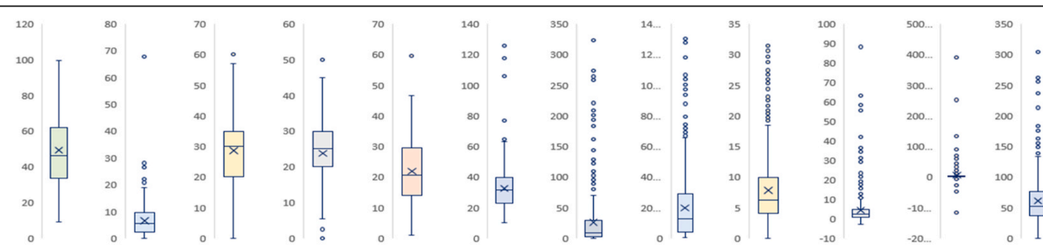
EPI differs from 9.3 to 99.3 with a mean value of 49.3 and the median value of 46.4. Also, the coefficient of variation CV statistics of EPI is relatively low. In general, CV values for each variable are at a high level, which means that many outliers exist in the variables. According to skewness, kurtosis statistics, and Kolmogorov-Smirnov test results, it can be said that all variables do not meet the normal distribution assumption. Likewise, TR, GE, POP, GDPPP, UN, INF, FDI, and PUBD have a right-skewed distribution, whereas CTR has a left-skewed distribution. All variables except EPI and ITR have a leptokurtic distribution.

4.2. Results of ML algorithms

In predicting the EPI and identifying the influential factors on the EPI with the greatest impact, a comprehensive analysis is conducted by employing a novel SL algorithm, which includes a total of six distinct ML algorithms. In the first step, the efficacy of each ML algorithm is assessed by using a set of goodness-of-fit metrics, and the obtained results are presented in Table 2, enabling a comparative analysis across the algorithms employed in this study.

Table 2
Performance metrics of ML algorithms.

ML Algorithm	Train Set			Test Set		
	R^2	RMSE	MAE	R^2	RMSE	MAE
RF	60.29%	13.1025	10.4838	80.22%	9.5086	8.0254
GBM	58.87%	13.1708	10.8981	80.21%	9.4715	8.0472
XGB	58.23%	13.3542	10.6012	73.53%	10.8141	9.0247
k-NN	46.35%	15.0982	12.3821	72.62%	10.9926	9.1731
CART	45.21%	15.4422	12.4519	52.67%	14.5269	12.1069
SVM	43.72%	18.9052	14.5926	67.58%	17.8521	14.3441



	EPI	TR	ITR	CTR	TB	GE	POP	GDPPP	UN	INF	FDI	PUBD
Minimum	9.3	0.0	0.0	0.0	1.0	10.1	0.1	651.9	0.0	-2.7	-115300.0	0.0
Q_{25}	33.6	2.5	20.0	20.0	13.9	23.0	2.5	4186.6	4.1	0.9	128.2	37.2
Median (Q_{50})	46.4	5.4	30.0	25.0	20.5	31.4	9.5	12553.9	6.2	2.4	873.0	52.3
Q_{75}	62.0	9.6	35.0	30.0	29.1	39.4	29.6	28933.3	10.0	4.8	3877.0	76.0
Maximum	99.3	67.7	60.0	50.0	59.6	125.7	327.4	130475.1	31.4	88.2	391104.0	304.1
Mean	49.3	6.5	28.7	23.8	21.8	32.6	26.4	19855.6	7.9	4.5	6801.4	61.1
Standard Deviation	20.8	5.5	13.1	8.9	10.3	13.3	46.0	21174.2	5.7	8.0	28068.7	37.9
Coefficient of Variation	42.2	84.0	45.8	37.3	47.2	40.9	174.3	106.6	72.8	180.2	412.7	62.1
Skewness	0.5	3.3	-0.2	-0.5	0.4	2.2	3.4	1.9	1.5	5.1	8.2	2.1
Kurtosis	-0.4	29.8	-0.3	0.8	-0.4	11.7	13.8	4.8	2.3	37.1	93.8	7.7
Kolmogorov-Smirnov	0.1	0.1	0.1	0.1	0.1	0.1	0.4	0.2	0.1	0.5	0.3	0.1
p value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Fig. 3. Preliminary statistics and Box-jenkins plots.

The 10-fold cross-validation approach (resampling) is used to find the optimal parameters of the models used and the best parameters are determined based on the RMSE, MAE, and R^2 statistics. Different values of mtry ranging from 2 to 12 are explored in the RF, and the optimal parameter value is determined as 7. In GBM, shrinkage is held constant at 0.1, interaction.depth and n.trees parameters range from 1 to 3, and from 50 to 150, respectively. The optimal values used for the model are n.trees = 50, interaction.depth = 3, shrinkage = 0.1. The XGB uses 5 different parameters: eta (from 0.3 to 0.4), max_depth (from 1 to 3), colsample_bytree (from 0.6 to 0.8), subsample (from 0.50 to 1.00), and nrounds (from 50 to 150). Optimum parameters are determined as eta = 0.3, max_depth = 2, colsample_bytree = 0.8, subsample = 1, and nrounds = 50. In k-NN, three parameters need to be optimized, which are kmax (from 5 to 9), distance (held constant at 2), and kernel (held at optimal). As a result, the optimum value of kmax is determined as 9. In CART, different values of cp ranging from 0.01 to 0.5 are examined and the optimal value is determined as 0.0673. Finally, in SVM, degree, scale, and C parameters ranging from 1 to 3, from 0.001 to 0.100, and from 0.25 to 1.00 are explored, respectively. The optimal values of these parameters are determined as degree = 1, scale = 0.001, and C = 0.5.

The RF algorithm demonstrates the highest R^2 value and the lowest RMSE and MAE statistics in the training set. Conversely, the SVM algorithm exhibits the lowest R^2 value and the highest RMSE and MAE statistics in the same set. These results indicate that the RF algorithm is the most effective in capturing the variation observed in the EPI. However, it is worth noting that the R^2 value of 60.29% obtained in the training set falls below the desired level for robust model performance.

In the second step, a novel SL algorithm is employed to enhance the model's predictive capabilities. The performance metrics of the SL algorithm are presented in Table 3.

Table 3 provides an overview of the model performance metrics for both the training and testing sets utilizing the SL algorithm. The results indicate that the RF, GBM, and XGB algorithms significantly contribute to the performance of the SL algorithm, whereas SVM has a minor impact and the CART and k-NN algorithms have no impact. More specifically, by calculating the predicted values of the SL algorithm as a weighted average of the predicted values from RF, GBM, XGB, and SVM, with corresponding weights of 51.7%, 30.6%, 17.1%, and 0.6%, respectively, it is evident that a robust and reliable model can be established. This weighted combination of algorithms results in notable improvements, as indicated by the substantial increase in the R^2 values from 60.29% achieved by the RF algorithm to 92.28% with the SL algorithm. Furthermore, the SL algorithm yields relatively lower RMSE and MAE statistics compared to each algorithm.

Additionally, in the testing set, the R^2, RMSE, and MAE values of the SL algorithm demonstrate acceptable levels of model interpretability and performance. To sum up, it is revealed that higher R^2 and lower RMSE and MAE statistics are reached by using the SL algorithm. In other words, the SL algorithm outperforms each algorithm.

4.3. Variable importance

Through the implementation of variable importance analysis, the relative influence of explanatory factors on the EPI is determined. Since the nature of the machine learning algorithms, variable importance analysis is used to determine the statistically significant variables affecting the response variable rather than the classical significance statistics (Mizumoto, 2023). This analysis enables the determination of the factors ranked in order of importance about their impacts on the EPI. Table 4 presents the variable importance results.

The weighted importance metrics represent the weighted average of the importance measures derived from each ML algorithm. In turn, the relative weighted importance is a measure that designates 100 as the most influential factor within the weighted importance metric. It is calculated by determining the relative importance measure accordingly.

Once the variable importance analysis outcomes are obtained separately with RF, XGB, GBM, and SVM algorithms, the order of importance can be differentiated. As a result, what is important is to interpret the variable importance according to the results of the SL algorithm, which is the best model. Based on the relative importance measures presented in Table 4, the factors with the greatest impact on the EPI are GDPPP, TR, TB, GE, and INF, respectively. These factors are deemed to play the most substantial role in the EPI. Conversely, some other factors, such as UN, POP, ITR, PUBD, FDI, and CTR exhibit statistically significant impacts on EPI, but their relative importance is quite lower. Consequently, policymakers should prioritize their efforts and focus on GDPPP, TR, TB, GE, and INF as key factors in their endeavors to enhance environmental performance.

Fig. 4 provides a visualization of the bivariate relationships between variables.

The x-axis implies the standardized values of the relevant factors, while the y-axis represents the predicted values obtained from the SL model for the EPI. It becomes evident that TB, GE, GDPPP, and FDI have the potential to positively influence EPI. Nevertheless, there exist significant thresholds, which are calculated by the SL algorithm, at which the impact of these factors on EPI changes considerably. These thresholds are determined as 11.5, 18.9, 19,855.5, and 6800.1 for TB, GE, GDPPP, and FDI, respectively. In contrast to the relationships observed with TR, ITR, and CTR, the relationships between UN versus EPI exhibit an "n-shaped" distribution. Conversely, EPI experiences a decrease as CTR increases.

Critical thresholds exist that alter the relationship between EPI and the respective factors. In other words, Thresholds find the values of the explanatory variables that change the response variable value at a statistically significant level and are calculated with the SL algorithm. More specifically, thresholds are obtained with the plotmo function in the R program. When the values of TR are lower than or equal to 13.7, a negative correlation is observed between EPI and TR. Conversely, while values exceed this threshold, a positive correlation is obvious. A similar relationship is observed between ITR and EPI, with a critical threshold of 30.1. Analyzing the relationship between CTR and EPI, it is apparent that when CTR values are lower than the average, the impact on EPI is not significant. However, while values surpass the average, EPI experiences a significant decrease as CTR values increase.

Finally, the visual analysis shows that the one-point increase (or decrease) in the explanatory variable causes more than (or less than) one-point change in the response variable, which means that there is a non-linear relationship between the response and explanatory variables

Table 3
Performance metrics of the SL algorithm.

SL	Risk	Coef.	Train Set			Test Set		
			R^2	RMSE	MAE	R^2	RMSE	MAE
CART	216.8	0.0%	92.28%	6.3942	5.1528	79.98%	9.6689	8.2842
RF	170.1	51.7%						
XGB	196.1	17.1%						
GBM	174.9	30.6%						
SVM	222.8	0.6%						
k-NN	258.2	0.0%						

Table 4
Relative and weighted importance Output.

Variable	SL		RF	XGB	GBM	SVM
	Weighted Relative Importance	**Weighted Importance**	51.7%	17.1%	30.6%	0.6%
GDPPP	**100.000**	83.476	100.000	100.000	46.971	50.423
TR	**39.822**	33.242	41.738	28.683	21.618	23.855
TB	**24.151**	20.161	23.310	27.470	10.702	22.842
GE	**11.437**	9.547	10.782	11.815	6.052	16.700
INF	**10.387**	8.671	9.181	10.857	4.795	100.040
UN	**9.713**	8.108	7.172	14.954	5.158	44.089
POP	**5.202**	4.342	5.434	3.255	2.919	13.854
ITR	**2.474**	2.066	3.614	0.090	0.090	25.689
PUBD	**2.384**	1.990	2.543	2.548	0.752	1.600
FDI	**2.298**	1.919	2.776	1.226	0.718	9.004
CTR	**0.602**	0.503	0.060	0.932	0.593	21.783

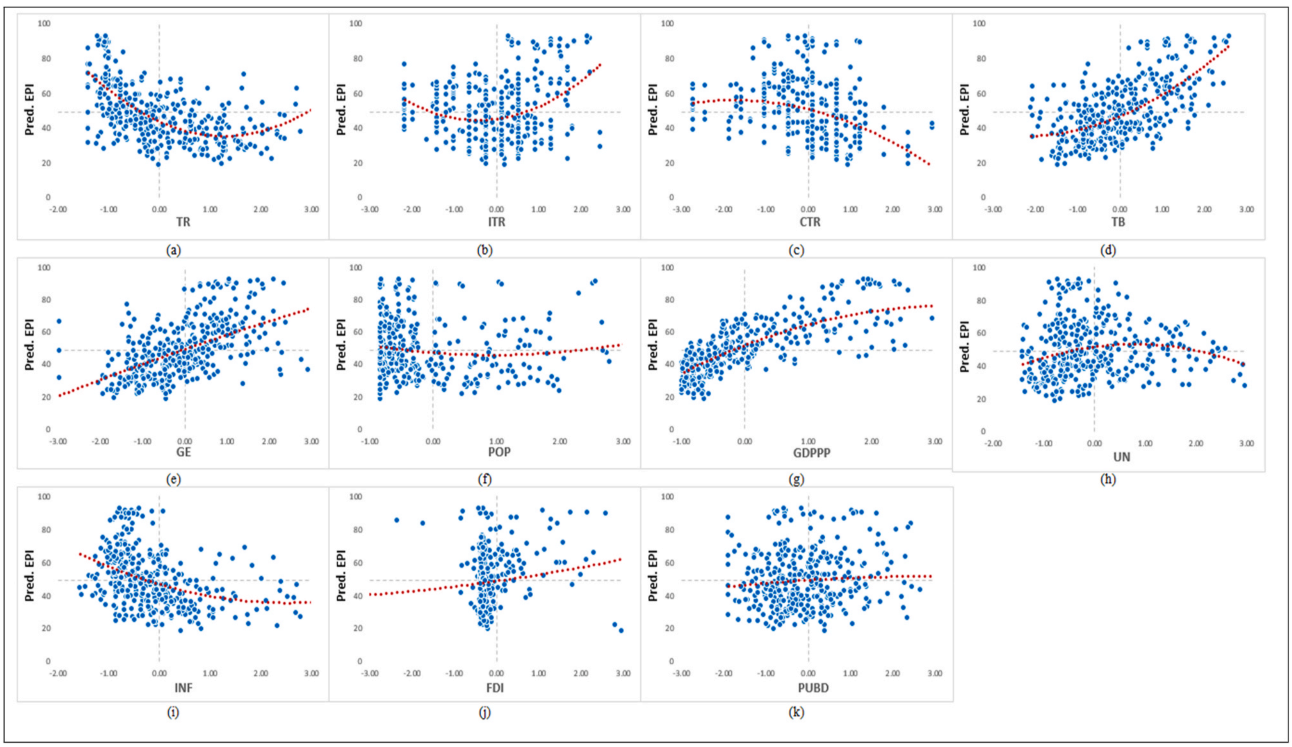


Fig. 4. Bivariate relationship between variables and significant thresholds.

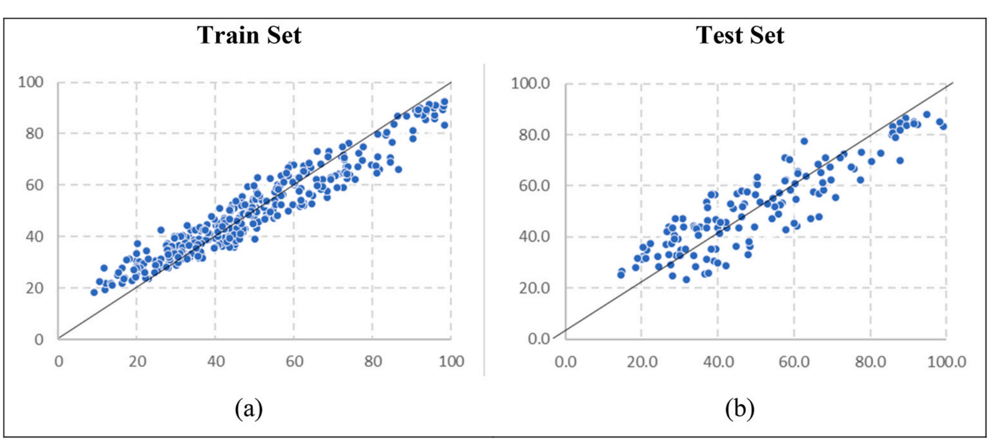


Fig. 5. Actual versus predicted values of SL algorithm.

as well. Also, the models used in the study are defined as non-linear models in ML approaches.

In the final analysis, Fig. 5 depicts the visualization of the actual values and predicted values derived from the SL algorithm.

The x-axis corresponds to the actual values, while the y-axis represents the predicted values. It is clear that all observations in both the training and testing datasets closely align with the 45-degree line. This alignment signifies proximity between the actual and predicted values. Therefore, it can be inferred that the model effectively captures and explains the variation observed in the EPI. In addition, actual versus predicted values scatter plots for each ML algorithm for train and test datasets are given in Figs. 6–7.

It can be revealed that the gap between actual versus predicted values obtained by the SL, RF, XGB, GBM, and k-NN is relatively small, while in the CART and SVM actual and predicted values are far from each other. The main reason for these results may be that the CART and SVM algorithms are based on binary split and hyperplane division, respectively; these algorithms may not effectively capture the complex relationship as in ensemble models such as RF, XGB, and GBM, even if parameter tuning is applied (Tatsat et al., 2020).

5. Conclusion, policy implications, and further research

5.1. Conclusion

By considering the increasing impact of the environment on humankind, the study aims to analyze the impact of a set of explanatory factors (e.g., tariff rate, income tax rate, corporate tax rate, tax burden, government expenditure, population, GDPPP, unemployment, inflation, FDI inflow, and public debt) on the EPI through the utilization of ML algorithms. In line with the aim of the study, a comprehensive dataset is compiled by incorporating information from two distinct data sources. Furthermore, a novel SL algorithm, which includes a total of six ML algorithms, is employed to effectively model the EPI and investigate its relationship with the aforementioned factors.

To address the issue of overfitting, the dataset is divided into two subsets: 70% for training samples and 30% for testing samples. Additionally, a 10-fold cross-validation technique is implemented for each ML model to establish a robust model. Initially, each ML algorithm is applied individually, and statistical metrics (e.g., R^2, RMSE, and MAE) are computed. The results indicate that the RF algorithm achieves better R^2, RMSE, and MAE values, while the SVM exhibits the poorest R^2, RMSE, and MAE statistics within the training set. However, the performance metrics of each model on the test set are insufficient for conclusive interpretation. Consequently, an ensemble model called the SL algorithm is employed to obtain a robust model. The SL algorithm surpasses the performance of all ML algorithms utilized. Furthermore, both the goodness-of-fit statistics for the train and test sets demonstrate significant accuracy. Notably, the predicted values of the SL algorithm are derived by computing a weighted average of the predicted values generated by the RF, GBM, XGB, and SVM models. The corresponding weights assigned to each model are 51.2%, 31.3%, 17.0%, and 0.7%, respectively.

5.2. Policy implications

According to the feature importance analysis, which shows the most influencing factors on the response variable, it is shown that the top five important factors that help policymakers to improve EPI are GDPPP, TR, TB, GE, and INF. In addition, it is also revealed that UN, POP, ITR, PUBD, FDI, and CTR have a significant impact on the EPI. Based on the study of Nie et al. (2022), it is revealed that economic and revenue growth hurt environmental performance metrics. Contrary to the results obtained by Nie et al. (2022) and Wang and Li (2021) show that the EPI metric decreases with the increase in GDPPP. In this study, similar to the study of Wang and Li (2021), it is shown that there is a positive correlation between GDPPP and EPI. On the other hand, this study shows that the higher level of GDP growth may cause a deterioration impact on the EPI. These results reveal that policymakers should consider economic growth to create a ready-to-action plan for improving environmental performance. The economic growth and GDPPP should be kept under control and measures should be taken to prevent uncontrolled growth.

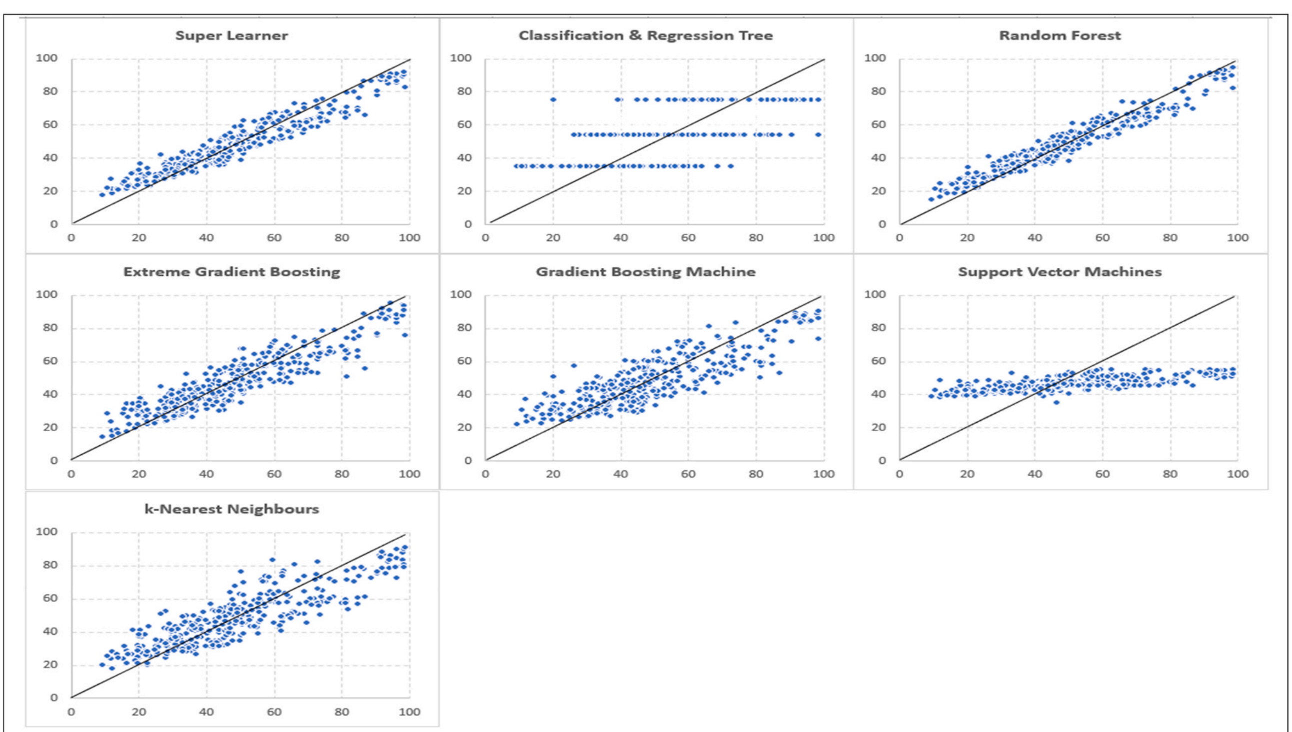


Fig. 6. Actual versus predicted values for train dataset.

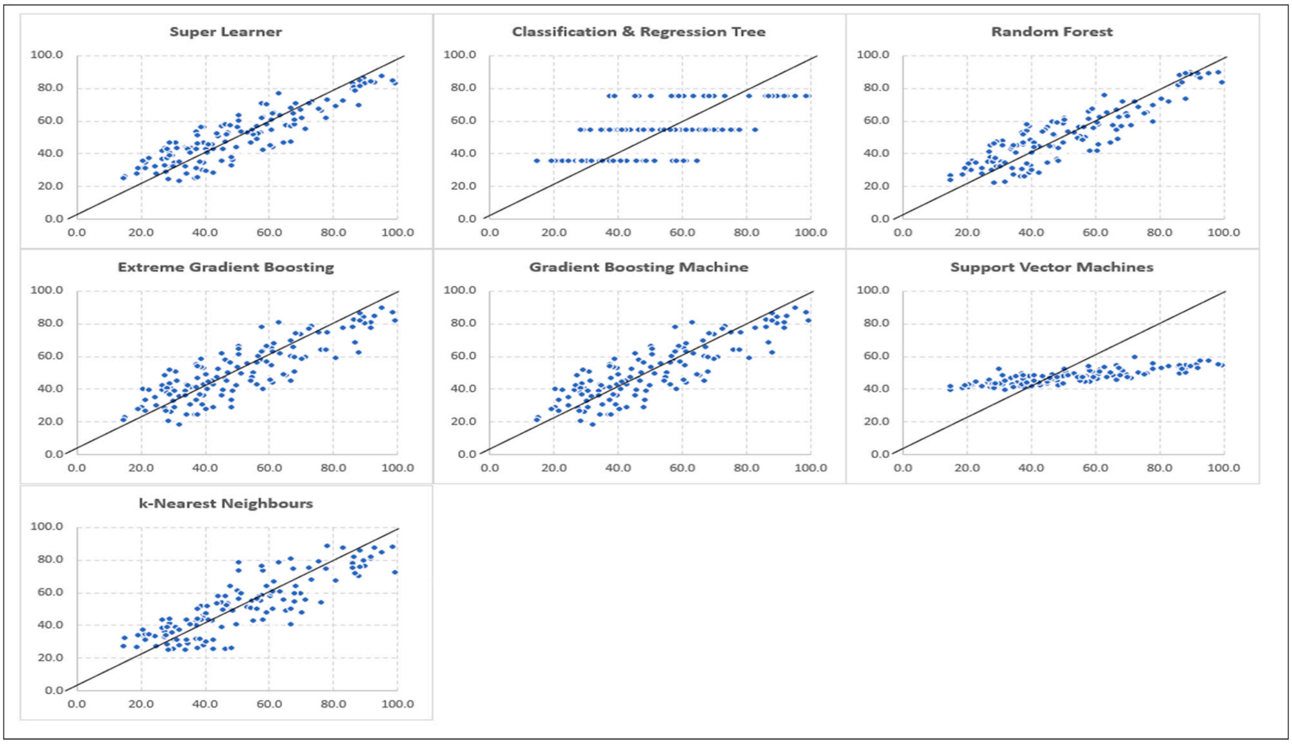


Fig. 7. Actual versus predicted values for test dataset.

The findings of this study highlight the tariff rate as the second most influential factor affecting the EPI. Consistent with existing literature, it is demonstrated that an increase in the tariff rate leads to a significant deterioration in the EPI. This observation aligns with the results emphasized in the study conducted by Kou et al. (2001), where it is shown that tariff limits and barriers exert a noteworthy negative impact on environmental performance metrics. Based on these results, policymakers should consider implementing tariff limits or cuts, especially for renewable energy sources and raw materials, to decrease CO_2 emissions and improve environmental performance. Implementation of measures such as tariff limits or cuts in these areas can contribute to improving environmental performance and fostering sustainable practices.

The tax burden, income tax, and corporate tax are frequently emphasized factors affecting environmental performance in the literature. Therefore, it is suggested that tax burden, income tax, and corporate tax rate should be fully- or partially exempted to encourage high technology-based innovation firms to produce, also it is shown that environmental tax has a great impact on decreasing CO_2 emissions as well (Ulucak et al., 2020; Esen et al., 2021; Wolde-Rufael and Mulat-Weldemeskel, 2021; Usman et al., 2022). The government should reconsider the aforementioned taxes to improve the EPI. In this study, it is shown that an increase in tax burden and income tax rate causes an increase in EPI while an increase in corporate tax rate causes a decrease in EPI. Thus, policymakers should consider decreasing corporate tax rates and improving tax burden and income tax to improve EPI.

In the extant literature, it is shown that government expenditure and inflation have a significant impact on environmental performance (Sadeh et al., 2020; Rahman et al., 2022). In line with the literature, this study shows that there is a positive correlation between government expenditure and EPI while the relationship between inflation and EPI is negative. Moreover, in this study, it is revealed that the impact of government expenditure on EPI has a linear structure. Also, especially investing in renewable energy and R&D cause an increase in EPI (Wolde-Rufael and Mulat-Weldemeskel, 2021; Dogan et al., 2022). On the other hand, a high level of inflation causes a low level of environmental performance. Thus, it is suggested that policymakers should focus on how to invest in renewable and green energy sources to improve environmental performance.

Also, policymakers should encourage foreign investors to increase the amount of foreign direct investments that can be used in R&D, renewable energy, and green energy sources. In addition, they should create action plans to decline inflation as well.

In the study of Xin et al. (2023), it is revealed that the impact of urbanization, unemployment rates, and population density versus per capita CO_2, which is used as an environmental quality metric is linear. Contrary to the study of Xin et al. (2023), this study shows that the relationship between population and unemployment versus EPI is significant and has a non-linear structure. Therefore, it can be suggested that policymakers should keep the unemployment rate at a minimum level. In addition, it is an issue that should be kept in mind the high unemployment rate hurts environmental metrics in the long run.

Similar to the relationship between unemployment versus EPI, the impact of population on EPI is non-linear. However, the impact of the population should be interpreted with the quality of waste management together (Jahn, 1998; UNDP, 2019; Nastase et al., 2019). Thus, it is suggested that policymakers should include a high-level waste management plan in the development plans of the country because low-level waste management causes a low-level environmental performance.

5.3. Limitations and further research

Although the study provides many policy recommendations, there are also some limitations. First, different ML or deep learning approaches can be included in the SL approaches to increase the model accuracy. Second, an up-to-date dataset can be used when it is shared with the public to compare the improvement of the model. Third, different environmental metrics, such as CO_2 emissions, ecological footprint, and load capacity factor, can be used as environmental proxies to bring different perspectives.

Disclosure statement

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics approval and consent to participate

Not applicable.

Consent for publication

The authors are willing to permit the Journal to publish the article.

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Mustafa Tevfik Kartal: Writing – review & editing, Writing – original draft, Conceptualization. **Özer Depren:** Writing – original draft, Validation, Software, Methodology, Formal analysis. **Serpil Kılıç Depren:** Writing – original draft, Conceptualization.

Declaration of competing interest

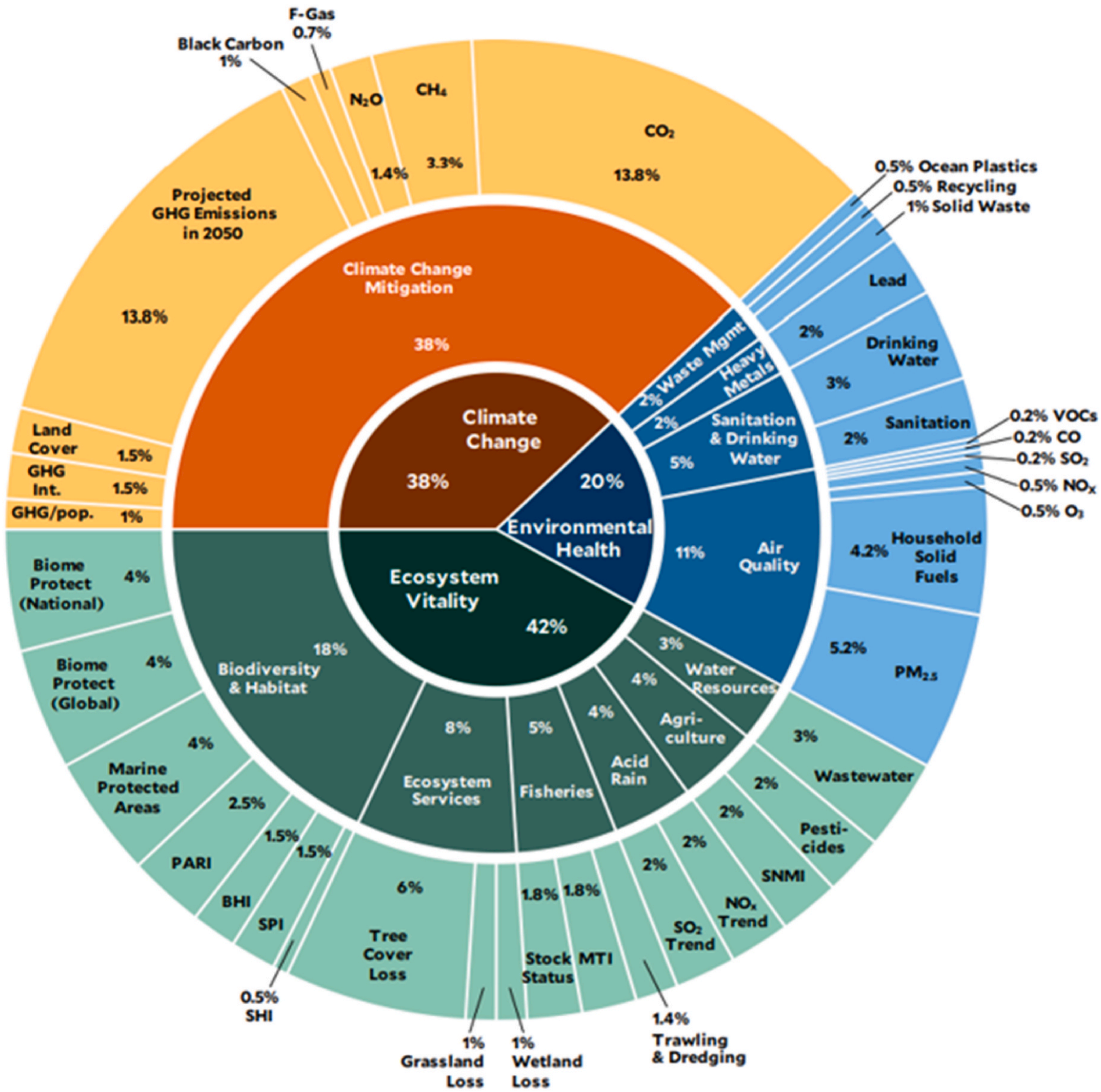
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Not applicable.

Appendix 1. Environmental Factors within Dimensions in EPI Methodology

Source: EPI 2022 Technical Report (Wolf et al., 2022)



Nomenclature

Abbreviations	
BRICS	Brazil, Russia, India, China, South Africa
CART	Classification and Regression Trees
CO_2	Carbon Dioxide
CV	Coefficient of Variation
EKC	Environmental Kuznets Curve
EPC	Environmental Phillips Curve
GBM	Gradient Boosting Machine
GDP	Gross Domestic Product
k-NN	k-Nearest Neighbors
MAE	Mean Absolute Error
ML	Machine Learning
R&D	Research and Development
R^2	Coefficient of Determination
RF	Random Forest
RMSE	Root Mean Square Error
SDGs	Sustainable Development Goals
SL	Super Learner

(*continued on next page*)

(*continued*)

SVM	Support Vector Machines
USD	United States Dollar
XGB	Extreme Gradient Boosting
Response Variable	
EPI	Environmental Performance Index
Explanatory Variables	
CTR	Corporate Tax Rate
GDPPP	Gross Domestic Product per Capita
FDI	Foreign Direct Investment
GE	Government Expenditure
INF	Inflation
ITR	Income Tax Rate
POP	Population
PUBD	Public Debt
TB	Tax Burden
TR	Tariff Rate
UN	Unemployment

References

Adekoya, O.B., Daniel, O.O., Ogunbowale, G.O., Al-Faryan, M.A.S., 2024. Unregulated economic activities and the environment: the role of fiscal policies. J. Environ. Manag. 349, 119533.

Ahmad, W., Ullah, S., Ozturk, I., Majeed, M.T., 2021. Does inflation instability affect environmental pollution? Fresh evidence from Asian economies. Energy Environ. 32 (7), 1275–1291.

Anser, M.K., Apergis, N., Syed, Q.R., Alola, A.A., 2021. Exploring a new perspective of sustainable development drive through environmental Phillips curve in the case of the BRICST countries. Environ. Sci. Pollut. Control Ser. 28 (35), 48112–48122.

Bekun, F.V., 2024. Race to carbon neutrality in South Africa: what role does environmental technological innovation play? Appl. Energy 354, 122212.

Chambers, R., 2011. Sustainable Livelihoods: an Opportunity for the. World Commission on Environment and Development. https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/873.

Chowdhury, T., Islam, S., 2017. Environmental Performance Index and GDP growth rate: evidence from BRICS countries. Environ. Econ. 8 (4), 31–36.

Clootens, N., 2017. Public debt, life expectancy, and the environment. Environ. Model. Assess. 22 (3), 267–278.

Dam, M.M., Kaya, F., Bekun, F.V., 2024. How does technological innovation affect the ecological footprint? Evidence from E-7 countries in the background of the SDGs. J. Clean. Prod. 443, 141020.

Dedeoğlu, M., Koçak, E., Uucak, Z.Ş., 2021. The impact of immigration on human capital and carbon dioxide emissions in the USA: an empirical investigation. Air Quality, Atmosphere & Health 14, 705–714.

Dogan, E., Hodžić, S., Šikić, T.F., 2022. A way forward in reducing carbon emissions in environmentally friendly countries: the role of Green Growth and environmental taxes. Economic Research-Ekonomska Istraživanja 35 (1), 5879–5894.

Esen, Ö., Yıldırım, D.Ç., Yıldırım, S., 2021. Pollute less or tax more? asymmetries in the EU environmental taxes-Eecological balance nexus. Environ. Impact Assess. Rev. 91, 106662.

Essandoh, O.K., Islam, M., Kakinaka, M., 2020. Linking international trade and foreign direct investment to CO_2 emissions: any differences between developed and developing countries? Sci. Total Environ. 712, 136437.

Fodha, M., Seegmuller, T., 2014. Environmental quality, public debt and economic development. Environ. Resour. Econ. 57 (4), 487–504.

Fu, Q., Chen, Y.E., Jang, C.L., Chang, C.P., 2020. The impact of international sanctions on environmental performance. Sci. Total Environ. 745, 141007.

Grossman, G.M., Krueger, A.B., 1991. Environmental Impacts of a North American Free Trade Agreement. National Bureau of Economic Research. No. 3914.

Hao, L.N., Umar, M., Khan, Z., Ali, W., 2021. Green growth and low carbon emission in G7 countries: how critical the network of environmental taxes, renewable energy and human capital is? Sci. Total Environ. 752, 141853.

Hao, Y., Wu, Y., Wu, H., Ren, S., 2020. How do FDI and technical innovation affect environmental quality? Evidence from China. Environ. Sci. Pollut. Control Ser. 27, 7835–7850.

Hashmi, R., Alam, K., 2019. Dynamic relationship among environmental regulation, innovation, CO_2 emissions, population, and economic growth in OECD countries: a panel investigation. J. Clean. Prod. 231, 1100–1109.

Hassan, S.T., Wang, P., Khan, I., Zhu, B., 2023. The impact of economic complexity, technology advancements, and nuclear energy consumption on the ecological footprint of the USA: towards circular economy initiatives. Gondwana Res. 113, 237–246.

Heritage, 2023. Data for Index of Economic Freedom. https://www.heritage.org/index.

Jahn, D., 1998. Environmental performance and policy regimes: explaining variations in 18 OECD-countries. Pol. Sci. 31 (2), 107–131.

Kartal, M.T., Pata, U.K., 2023. Impacts of renewable energy, trade globalization, and technological innovation on environmental development in China: evidence from various environmental indicators and novel quantile methods. Environmental Development, 100923.

Kartal, M.T., Kılıç Depren, S., Kirikkaleli, D., 2023a. Asymmetric effect of political stability on production-based CO_2 emissions in the UK: long-run evidence from nonlinear ARDL and frequency domain causality. Environ. Sci. Pollut. Control Ser. 30 (12), 33886–33897.

Kartal, M.T., Pata, U.K., Destek, M.A., Caglar, A.E., 2023b. Environmental effect of clean energy research and development investments: evidence from Japan by using load capacity factor. J. Clean. Prod. 416, 137972.

Kashem, M.A., Rahman, M.M., 2020. Environmental phillips curve: OECD and asian NICs perspective. Environ. Sci. Pollut. Control Ser. 27 (25), 31153–31170.

Khan, I., Hou, F., Le, H.P., 2021a. The impact of natural resources, energy consumption, and population growth on environmental quality: fresh evidence from the United States of America. Sci. Total Environ. 754, 142222.

Khan, M.K., Abbas, F., Godil, D.I., Sharif, A., Ahmed, Z., Anser, M.K., 2021b. Moving towards sustainability: how do natural resources, financial development, and economic growth interact with the ecological footprint in Malaysia? A dynamic ARDL approach. Environ. Sci. Pollut. Control Ser. 28 (39), 55579–55591.

Khan, S.A.R., Zhang, Y., Kumar, A., Zavadskas, E., Streimikiene, D., 2020. Measuring the impact of renewable energy, public health expenditure, logistics, and environmental performance on sustainable economic growth. Sustain. Dev. 28 (4), 833–843.

Kılıç Depren, S., Depren, Ö., 2021. How economic freedom affect the ease of doing business? Cross-country quantile regression mixture approach. Sigma J. Eng. Nat. Sci. 39 (4), 434–442.

Kılıç Depren, S., Kartal, M.T., Çoban Çelikdemir, N., Depren, Ö., 2022. Energy consumption and environmental degradation nexus: a systematic review and meta-analysis of fossil fuel and renewable energy consumption. Ecol. Inf. 70, 101747.

Koçak, E., Ulucak, Z.Ş., 2019. The effect of energy R&D expenditures on CO_2 emission reduction: estimation of the STIRPAT model for OECD countries. Environ. Sci. Pollut. Control Ser. 26, 14328–14338.

Kou, H.I., Hu, J.L., Hwang, H., 2001. Tariff policy and environmental qualities of imported goods. J. Econ. Integrat. 313–343.

Kumar, S., Giridhar, V., Sadarangani, P., 2019. A cross-national study of environmental performance and culture: implications of the findings and strategies. Global Bus. Rev. 20 (4), 1051–1068.

Le, H.N.D., Hoang, D.M., 2022. Economic sanctions and environmental performance: the moderating roles of financial market development and institutional quality. Environ. Sci. Pollut. Control Ser. 29 (13), 19657–19678.

Li, H., Xu, X., Rafique, M.Z., Naqvi, S.A.A., Nadeem, A.M., 2021. What drives the environmental performance? Evidence from asian countries. Pol. J. Environ. Stud. 30 (4), 3685.

Li, Z., Huang, Z., Failler, P., 2019. Impact of foreign direct investment on environmental performance. Sustainability 11 (13), 3538.

Lin, B., Ullah, S., 2024. Evaluating forest depletion and structural change effects on environmental sustainability in Pakistan: through the lens of the load capacity factor. J. Environ. Manag. 353, 120174.

Liu, Y., Li, B., Yang, S., Li, Z., 2024. Handling missing values and imbalanced classes in machine learning to predict consumer preference: demonstrations and comparisons to prominent methods. Expert Syst. Appl. 237, 121694.

Mirzaei, A., Carter, S.R., Patanwala, A.E., Schneider, C.R., 2022. Missing data in surveys: key concepts, approaches, and applications. Res. Soc. Adm. Pharm. 18 (2), 2308–2316.

Mizumoto, A., 2023. Calculating the relative importance of multiple regression predictor variables using dominance analysis and random forests. Lang. Learn. 73 (1), 161–196.

Nastase, C., Chaşovschi, C.E., State, M., Scutariuf, A.L., 2019. Municipal waste management in Romania in the context of the EU. A stakeholders' perspective. Technol. Econ. Dev. Econ. https://doi.org/10.3846/tede.2019.10295.

Ng, C., Yii, K.J., Lau, L., Go, Y., 2022. Unemployment rate, clean energy, and ecological footprint in OECD countries. Environ. Sci. Pollut. Control Ser. 30 (15), 42863–42872.

Nie, G.Q., Zhu, Y.F., Wu, W.P., Xie, W.H., Wu, K.X., 2022. Impact of voluntary environmental regulation on green technological innovation: evidence from Chinese manufacturing enterprises. Front. Energy Res. 10, 889037.

Pata, U.K., Alola, A.A., Erdogan, S., Kartal, M.T., 2023a. The influence of income, economic policy uncertainty, geopolitical risk, and urbanization on renewable energy investments in G7 countries. Energy Econ. 128, 107172.

Pata, U.K., Kartal, M.T., Erdogan, S., Sarkodie, S.A., 2023b. The role of renewable and nuclear energy R&D expenditures and income on environmental quality in Germany: scrutinizing the EKC and LCC hypotheses with smooth structural changes. Appl. Energy 342, 121138.

Pimonenko, T.V., Liulov, O.V., Chygryn, O.Y., 2018. Environmental Performance Index: relation between social and economic welfare of the countries. Environ. Econ. 9 (3), 1–11.

Rahman, S.U., Faisal, F., Sami, F., Ali, A., Chander, R., Amin, M.Y., 2022. Investigating the nexus between inflation, financial development, and carbon emission: empirical evidence from FARDL and frequency domain approach. Journal of the Knowledge Economy. https://doi.org/10.1007/s13132-022-01076-w.

Rajput, D., Wang, W.J., Chen, C.C., 2023. Evaluation of a decided sample size in machine learning applications. BMC Bioinf. 24 (1), 48.

Ramezan, C.A., Warner, T.A., Maxwell, A.E., Price, B.S., 2021. Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. Rem. Sens. 13 (3), 368.

Raza, A., Sui, H., Jermsittiparsert, K., Żukiewicz-Sobczak, W., Sobczak, P., 2021. Trade liberalization and environmental performance index: mediation role of climate change performance and greenfield investment. Sustainability 13 (17), 9734.

Rehman, M.A., Sabir, S.A., Bukhari, A.A.A., Sharif, A., 2023. Do globalization and human capital an opportunity or threat to environmental sustainability? Evidence from emerging countries. J. Clean. Prod. 418, 138028.

Sadeh, A., Radu, C.M., Feniser, C., Borşa, A., 2020. Governmental intervention and its impact on growth, economic development, and technology in OECD countries. Sustainability 13 (1), 166.

Saidi, K., Omri, A., 2020. Reducing CO_2 emissions in OECD countries: do renewable and nuclear energy matter? Prog. Nucl. Energy 126, 103425.

Sala, O.E., Stuart Chapin, F., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., 2000. Global biodiversity scenarios for the year 2100. Science 287 (5459), 1770–1774.

Shang, Y., Razzaq, A., Chupradit, S., An, N.B., Abdul-Samad, Z., 2022. The role of renewable energy consumption and health expenditures in improving load capacity factor in ASEAN countries: exploring new paradigm using advance panel models. Renew. Energy 191, 715–722.

Sharif, A., Kartal, M.T., Bekun, F.V., Pata, U.K., Foon, C.L., Kılıç Depren, S., 2023. Role of green technology, environmental taxes, and green energy towards sustainable environment: insights from sovereign Nordic countries by CS-ARDL approach. Gondwana Res. 117, 194–206.

Shittu, W., Adedoyin, F.F., Shah, M.I., Musibau, H.O., 2021. An investigation of the nexus between natural resources, environmental performance, energy security and environmental degradation: evidence from Asia. Resour. Pol. 73, 102227.

Stoian, M., Brad, L., Zaharia, A., 2022. Drivers of the European union's environmental performance. Front. Environ. Sci. 10.

Tatsat, H., Puri, H., Lookabaugh, B., 2020. Machine Learning and Data Science Blueprints for Finance. O'Reilly Media.

Ullah, S., Apergis, N., Usman, A., Chishti, M.Z., 2020. Asymmetric effects of inflation instability and GDP growth volatility on environmental quality in Pakistan. Environ. Sci. Pollut. Control Ser. 27, 31892–31904.

Ullah, S., Luo, R., Adebayo, T.S., Kartal, M.T., 2023. Paving the ways toward sustainable development: the asymmetric effect of economic complexity, renewable electricity, and foreign direct investment on the environmental sustainability in BRICS-T. Environ. Dev. Sustain. 1–25. https://doi.org/10.1007/s10668-023-03085-4.

Ulman, S.R., Mihai, C., Cautisanu, C., Brumă, I.S., Coca, O., Stefan, G., 2021. Environmental performance in EU countries from the perspective of its relation to human and economic wellbeing. Int. J. Environ. Res. Publ. Health 18 (23), 12733.

Ulucak, R., Danish, K., Kassouri, Y., 2020. An assessment of the Environmental Sustainability Corridor: investigating the non-linear effects of environmental taxation on CO_2 emissions. Sustain. Dev. 28 (4), 1010–1018.

Ulussever, T., Kartal, M.T., Kılıç Depren, S., 2023. Effect of income, energy consumption, energy prices, political stability, and geopolitical risk on the environment: evidence from GCC countries by novel quantile-based methods. Energy Environ. https://doi.org/10.1177/0958305X231190351.

UNDP, 2019. Human Development Report 2019: beyond Income, beyond Averages, beyond Today: Inequalities in Human Development in the 21st Century (New York).

United Nations, 2015. Transforming our world: the 2030 agenda for sustainable development. https://sdgs.un.org/2030agenda.

Usman, M., Jahanger, A., Radulescu, M., Balsalobre-Lorente, D., 2022. Do nuclear energy, renewable energy, and environmental-related technologies asymmetrically reduce ecological footprint? Evidence from Pakistan. Energies 15 (9), 3448.

Van der Laan, M.J., Polley, E.C., Hubbard, A.E., 2007. Super learner. Stat. Appl. Genet. Mol. Biol. 6 (1) https://doi.org/10.2202/1544-6115.1309.

Van der Laan, M.J., Rose, S., 2011. Targeted Learning: Causal Inference for Observational and Experimental Data, vol. 4. Springer, New York.

Wang, Q., Li, L., 2021. The effects of population aging, life expectancy, unemployment rate, population density, per capita GDP, urbanization on per capita carbon emissions. Sustain. Prod. Consum. 28, 760–774.

Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res. 30 (1), 79–82.

Wolde-Rufael, Y., Mulat-Weldemeskel, E., 2021. Do environmental taxes and environmental stringency policies reduce CO_2 emissions? Evidence from 7 emerging economies. Environ. Sci. Pollut. Control Ser. 28, 22392–22408.

Wolf, M.J., Emerson, J.W., Esty, D.C., de Sherbinin, A., Wendling, Z.A., et al., 2022. 2022 Environmental Performance Index. Yale Center for Environmental Law & Policy, New Haven, CT.

Xie, Q., Yan, Y., Wang, X., 2023. Assessing the role of foreign direct investment in environmental sustainability: a spatial semiparametric panel approach. Econ. Change Restruct. 56 (2), 1263–1295.

Xin, Y., Yang, S., Rasheed, M., 2023. Exploring the impacts of education and unemployment on CO_2 emissions. Ekonomska Istrazivanja-Economic Research 36 (2), 2110139.

Yale Center for Environmental Law & Policy, 2023. Data of environmental performance index. https://epi.yale.edu/downloads.

Yang, X., Li, N., Mu, H., Ahmad, M., Meng, X., 2022. Population aging, renewable energy budgets and environmental sustainability: does health expenditures matter? Gondwana Res. 106, 303–314.

Yuelan, P., Akbar, M.W., Zia, Z., Arshad, M.I., 2022. Exploring the nexus between tax revenues, government expenditures, and climate change: empirical evidence from Belt and Road Initiative countries. Econ. Change Restruct. 55 (3), 1365–1395.

Zhang, Q., Zhang, S., Ding, Z., Hao, Y., 2017. Does government expenditure affect environmental quality? Empirical evidence using Chinese city-level data. J. Clean. Prod. 161, 143–152.