Journal of Capital Markets Studies

89

Received 16 September 2025
Revised 15 October 2025
Accepted 8 November 2025

Machine learning methods for investigating the nexus between FDI and environmental degradation: evidence from Sub-Saharan African countries

Nuray Tezcan
Department of Management Information Systems, Haliç University, Istanbul, Turkey

Abstract

Purpose – The purpose of the study is to investigate whether foreign direct investment (FDI) has an impact on environmental degradation in Sub-Saharan African (SSA) countries using machine learning (ML) methods for the years between 2002 and 2021.

Design/methodology/approach – In this study, k-nearest neighbour (k-NN), support vector machine (SVM) and random forest (RF) machine learning algorithms were used, and their performances were compared based on root mean squared error (RMSE), *R*-squared (R2), mean absolute error (MAE) criteria, respectively. While carbon dioxide (CO_2) emissions are used as an indicator of environmental degradation, in addition to FDI, gross domestic product (GDP) per capita, urbanization, renewable energy consumption, trade openness, population density, natural resources rents, governance and inflation are used as explanatory variables in the study to obtain a comprehensive perspective. The dataset consists of 46 countries and is compiled from the World Development Indicators Database, World Bank.

Findings – Among ML methods, it is found that RF has the best performance based on the performance evaluation criteria. It is found that there is no evidence that FDI has an impact on environmental degradation in SSA countries and GDP per capita, renewable energy consumption and urbanization are the most important features affecting carbon dioxide emissions. According to the findings, renewable energy consumption positively affects environmental degradation, whereas GDP per capita and urbanization negatively affect.

Originality/value – In the literature, while the relationship between environmental degradation and FDI is examined through econometric analyses, there are very few studies using machine learning method to investigate this relationship. Therefore, this study is an attempt to fill this gap in the literature, and it provides valuable insights into the applicability of ML methods.

Keywords Sub-Saharan African countries, Environmental degradation, Machine learning, Foreign direct investment

Paper type Research article

1. Introduction

Environmental sustainability has gained great importance after the sustainable development concept that came to the agenda with the publication of the Brundlant Report in 1987 and the Sustainable Development Goals (SDGs) declared in 2015. Especially after the Industrial Revolution, greenhouse gas (GHG) emissions, which largely consist of carbon dioxide (CO_2), have increased, leading to global warming and climate change. The doubling of global GHG emissions between 1970 and 2022 and the increase of over 60% in global CO_2 emissions between 1990 and 2022 are examples of the negative changes that have occurred (Statista, 2024). Moreover, temperatures continue to rise. According to a report published by the United Nations (UN), 2024 was the hottest year on record, approximately 1.55°C above pre-industrial levels (UN, 2025).

Journal of Capital Markets Studies
Vol. 10 No. 1, 2026
pp. 89-103
Emerald Publishing Limited
2514-4774
DOI 10.1108/JCMS-09-2025-0136

While the negative effects of global warming are seen around the world, when Sub-Saharan African (SSA) countries [1] are examined, it can be seen that the situation is not very different from other regions of the world. SSA countries account for about 15% of the world population and are mostly low-income or lower-middle income countries. Depending on the income level, these countries are not industrialized; therefore, the contribution of SSA countries in terms of CO_2 emissions is lower than the other regions in the world. However, CO_2 emissions continue to increase as in the rest of the world. Moreover, the rate of surface temperature increase in the African continent has been generally higher than the global average. South Africa, Nigeria, Angola, Sudan and Kenya are the largest CO_2 emitters in SSA countries and GHG emissions are mostly caused by changes in land use, forestry and agriculture in African countries (Statista, 2025). Figure 1 presents CO_2 emissions released by SSA countries for the period 2000 and 2020.

On the other hand, foreign direct investment (FDI) is an important indicator for the countries in terms of economic growth. There is an argument in the literature about the advantages and disadvantages of the FDI, and while some studies have proved that this leads to environmental degradation, other studies have found the opposite. Accordingly, there are two hypotheses in the literature regarding the impact of the FDI on environmental degradation: The Pollution Haven Hypothesis and the Pollution Halo Hypothesis. According to the Pollution Haven Hypothesis, intensive polluting sectors operating in developed countries are shifted through FDI to developing countries where environmental sustainability improvements are not strict. Thus, these developing countries transform pollution havens. On the other hand, Pollution Halo Hypothesis states that developed countries transfer their pollution-reducing technologies, renewable energy-related technologies and energy-conserving technologies to host countries through FDI. Thus, they lead to a decline in environmental degradation (Duan and Jiang, 2021).

With the increasing importance of environmental sustainability, it has become inevitable to take some measures. Reducing the use of fossil fuels and encouraging and expanding the use of renewable energy sources are among the priorities of the countries. Another indicator that needs to be reduced for environmental sustainability is total natural resources rents (TNRR) that is assumed to be a blessing or a curse according to the level of democracy and governance of countries and can cause environmental degradation (Ploeg, 2011). The adverse effect of TNRR can be observed through the Climate Change Performance Index (CCPI), which is one of the most important indices regarding environmental sustainability. In this index, countries are divided into five groups based on their overall scores: very high, high, medium, low and very low. When examining countries that consistently ranked in the “very low” group during the 2018–2023 period, it was observed that most of these countries are global leaders in oil, gas and coal production and also have very high TNRR values (Burck *et al.*, 2023).

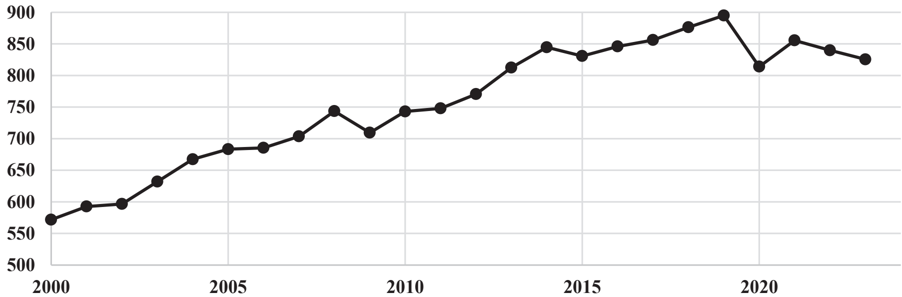

Figure 1. CO_2 emissions (total) excluding LULUCF (Mt CO_2e) of SSA Countries between 2000 and 2023.
Source: WDI, WB

When the SSA countries are examined in terms of these indicators for the last two decades in Figure 2, it is seen that there is no significant change in terms of fossil fuel energy consumption, renewable energy consumption. On the other hand, it shows that renewable energy consumption level is more than about 70% and this level is very high in comparison with other regions (Espoir *et al.*, 2023). For the TNRR indicator, it is seen that there are some fluctuations; accordingly, the TNRR value decreased between 2011 and 2015, then remained almost at the same level after 2015.

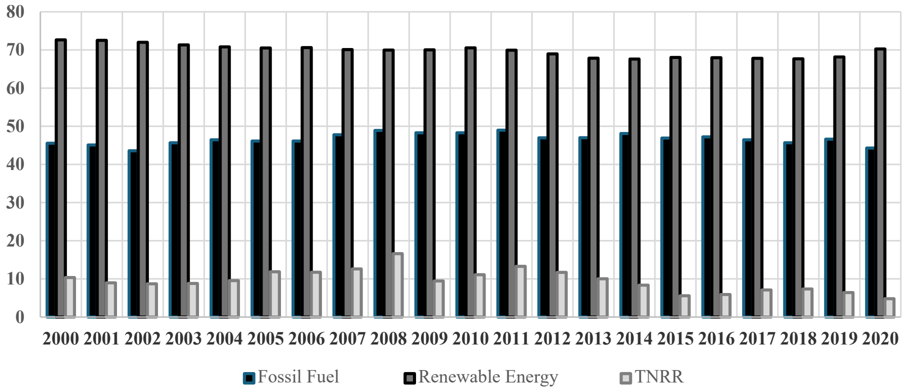

Figure 2. Fossil fuel energy consumption (%), renewable energy consumption (%) and total natural resources rents (%) for SSA countries from 2000 to 2020. Source: WDI, WB

The following parts of this study are organized as follows: The second part presents the comprehensive literature review and the third part presents the dataset, features and the methods used. In the following part, results obtained from the analyses are provided, and the last part is dedicated to discussion and conclusion.

2. Literature review

In the literature, numerous empirical studies were conducted regarding the relationship between FDI and environmental degradation for different countries for different time spans. In these studies, several factors affecting environmental degradation were employed, such as urbanization, TNRR (Adams and Klobodu, 2017; Dua and Xiab, 2018; Koirala and Pradhan, 2020; Sanchez and Ortega, 2020; Chien *et al.*, 2023; Alhassan and Kwakwa, 2023), renewable energy (Güney, 2019) and governance (Bokpin, 2017; Safdar *et al.*, 2022; Güney, 2017, 2022).

With respect to SSA countries, several studies examining relationship between FDI and environmental degradation have been conducted and these studies employed econometrics analyses. Some of these are stated as follows:

Wang and Dong (2019) investigated drives of environmental degradation in SSA countries for the period 1990–2014. In this study, ecological footprint as indicator of environmental degradation, renewable and non-renewable energy consumption, GDP and urban population variables were used. The results showed that non-renewable energy consumption, GDP and urban population positively affect environmental degradation, whereas renewable energy negatively affects. In another study conducted by Dhrifi *et al.* (2020), causality between FDI, CO_2 and poverty was studied based on three regions: Asia, Africa and Latin America and total of 98 developing countries. Moreover, GDP per capita, urbanization, population, financial development, inflation, education, trade openness, infrastructure and energy consumption variables were used in the simultaneous equations. It was found that there is unidirectional causal relationship that runs from FDI to CO_2 emission for the whole sample and, additionally, a positive impact of FDI on CO_2 emission in African countries.

Unlike other studies testing environmental Kuznets curve (EKC) hypothesis, Espoir *et al.* (2022) investigated the impact of CO_2 emissions and temperature on income level utilizing 47 African countries between the years 1995 and 2016. Findings showed that increasing in average temperature reduces income, while increasing in CO_2 emissions drives income level. Essandoh *et al.* (2020) focused on relationships among CO_2 emissions, international trade and FDI inflows, real GDP per capita, renewable energy consumption and primary energy consumption using the panel pooled mean group-autoregressive distributive lag method. In this study, 52 developed and developing countries were analyzed for the period 1991–2014 and compared. According to the results, there is a positive relationship between CO_2 emissions and FDI in developing countries, whereas there is a negative long-run relationship between CO_2 emissions and trade openness in developed countries. This result supported the insight that there is a transfer of high emission-intensive production units from developed countries to developing countries.

Baloch *et al.* (2020) studied the impact of income inequality and poverty on CO_2 emissions for the 40 Sub-Saharan African countries between 2010 and 2016. GDP per capita, inflation, population, economic freedom and access to electricity variables were included in the study, and Driscoll–Kraay regression estimator was used. According to findings, inflation, economic freedom and access to electricity have no impact on CO_2 emissions whereas GDP per capita, income inequality and poverty have a positive impact and population has negative.

Karim *et al.* (2022) investigated the impact of institutional quality on CO_2 emissions in 30 SSA countries for the period from 2000 to 2021 using on EKC model. For the institutional quality, six dimensions from the World Governance Indicators (WGI) were used and cross section autoregression distributed lag was employed as the method. Additionally, economic growth, industrialization, energy consumption and population growth were used. Findings indicated that corruption control, regulatory quality and the rule of law reduce CO_2 emissions. Also, it is found that there is positive relationship between economic growth and CO_2 emissions.

Espoir and Oyadeyi (2025) investigated whether the Pollution Haven Hypothesis is valid within African countries for the period 1970–2022. In the study using ecological footprint as an indicator of environmental degradation, top five countries in Africa in terms of FDI were focused on, and the Fourier Seemingly Unrelated Regressions Mean Group estimator was used. Findings pointed out that FDI significantly affects environmental degradation in Africa and the Pollution Haven Hypothesis was validated.

Akadiri *et al.* (2024) investigated the impact of financial globalization and natural resources rents on CO_2 emissions in Nigeria between 1970 and 2020 using Granger causality analysis. They found that there is a one-way causal effect running from financial globalization and natural resource rents to carbon emissions. In another study conducted by Prempeh *et al.* (2024), it is found that natural resource rents have had a negligible negative impact on carbon emissions in the short term but have shown a significant positive impact in the long term in Ghana.

Sakariyahu *et al.* (2024) studied the relationship between quality of life and environmental degradation for 31 African countries for the period 2000–2018 using Lewbel two stages least squares method. Renewable energy, inflation, FDI, GDP, trade openness, governance, population growth were used as control variables. They found that there is a negative relationship between quality of life and environmental degradation. Also, it has been seen that quality of life and FDI are negatively correlated.

To give examples from other countries, Huang *et al.* (2025) have found that FDI and trade openness caused environmental pollution in Pakistan, and Kutlu and Atis (2024) studied the countries with the highest CO_2 emissions, and the results showed that there is a U-shaped relationship between FDI and carbon emissions for the period 1996 and 2022.

In addition to the econometric analyses, more recently, there have been conducted studies using machine learning (ML) methods to determine factors affecting environmental degradation or environmental sustainability. In a review study conducted by Li *et al.* (2025), 105 studies regarding maritime environmental sustainability were examined in terms of ML algorithms used in these studies, and it has been seen that gradient boosting machines, artificial neural networks and random forest (RF) algorithms are frequently used in the literature related to this area. In another study related to environmental sustainability, Asrol *et al*. (2021) used support vector machine (SVM) as ML algorithm to assess the sustainability performance of the bioenergy industry and predicted the test data with 98.32% accuracy. Wang *et al*. (2024) investigated influencing factors of urban low-carbon innovation using both single and ensemble ML methods. In this study, ten factors, such as expenditure on research, carbon emissions, local general budgetary revenue and education expenditure, were determined as the most important among 41 factors. Also, it has found that LightGBM has the best performance in terms of prediction.

Imam *et al*. (2024) predicted the air quality in two different regions of India using several ML methods. In this study, the SVM model obtained the best accuracy of 97.98% on the first dataset, while the RF model obtained the highest accuracy of 93.29% on the second dataset.

Osman *et al*. (2025) explored the relationship between climate change and economic growth in Somalia using random forest regression and Bayesian regression. They found that temperature and CO_2 emissions negatively influence GDP, while rainfall has a positive effect. Additionally, it is found that random forest regression has the better than Bayesian regression in terms of prediction performance.

Khan *et al*. (2024), investigated impact of green investment on environmental degradation in developing countries using machine learning-based methods for the period 2009–2019. Findings indicated that green investment significantly reduces environmental degradation.

In another study used machine learning method, Espoir *et al*. (2024) investigated key determinants of ecological footprint in 39 African countries for the period 1996–2018. In the first step of the analysis, the least absolute shrinkage and selection operator (LASSO) was used to reveal the most important features among 26 features. Then, partialing-out LASSO instrumental variable regression and Bayesian model averaging techniques were employed to obtain marginal effects of the variables. GDP per capita and its square, domestic credit to the private sector, the share of total natural resources, population density, renewable energy, the share of manufacturing value added in GDP, human capital, gross fixed capital formation, poverty incidence, voice accountability, government effectiveness and rule of law are found significant features affecting ecological footprint.

Agan *et al*. (2025) investigated factors affecting CO_2 emissions in SSA countries using different ML algorithms for the period 1990–2022. Energy efficiency, government stability, clean energy, GDP per capita and population were used in this study, RF algorithm has the best performance and population, clean energy and GDP per capita were found to be the most important features affecting CO_2 emissions.

3. Research design

3.1 Data and features used

To determine the impact of FDI on environmental degradation in SSA countries, the World Development Indicators (WDI) and the WGI databases obtained from the World Bank (WB) were employed for the period 2002–2021 (WB, 2025a, b).

In the analyses CO_2 emissions were used as an indicator of environmental degradation while the features were determined in line with the studies given in the literature review. The list of the features, their abbreviations and sources are given in Table 1 below.

For the governance feature, the mean value was calculated for the six dimensions of institutional quality, which are voice and accountability, political stability and absence of violence/terrorism, government effectiveness, regulatory quality, rule of law and control of corruption.

3.2 Data preprocessing

Although there are 48 SSA countries according to the WB classification, the sample consisted of 46 countries and the time span was determined as 2002–2021 due to data availability. Before the analysis, dataset was examined in terms of missing values and the existence of outliers. Since the missing value percentages of the features were at an acceptable level (maximum

Table 1. List of the features used in ML methods

Target variable	Abbreviation	Source
Carbon dioxide (CO_2) emissions excluding LULUCF per capita (t CO_2e/capita)	CO_2	WDI
Features		
Foreign direct investment, net inflows (% of GDP)	FDI	WDI
Renewable energy consumption (% of total final energy consumption)	REnergy	WDI
GDP per capita, PPP (current international $)	GDP	WDI
Trade openness (% of GDP)	Trade	WDI
Inflation, consumer prices (annual % growth)	Inflation	WDI
Population density (people per sq. km of land area)	Pop	WDI
Urban population (% of total population)	Urban	WDI
Total natural resources rents (% of GDP)	TNRR	WDI
Governance	Gov	WGI

Source(s): Author's own work

level is 13%), the missing values of each feature were imputed with the mean value of that feature. To remove skewness of the features was taken their natural logarithm and outliers were determined box-plots and standardization. After, highly extreme values were dropped from the dataset; therefore, the final dataset consisted of 876 country-years observations. Data analysis will be performed based on the R programming language and R Studio. Before the analyses, the data is divided into 80 and 20% as training and test data.

3.3 Methods used

Unlike other studies conducted previously, ML methods were employed in this study to reveal the nexus between variables. ML is programming computers to optimize a performance criterion using example data or past experience and uses the theory of statistics in building mathematical models, because the core task is making inferences from a sample (Alpaydın, 2020). Machine learning aims to generate an unknown rule from instances or generalize a detectable pattern.

In this study, three supervised ML methods, k-nearest neighbour (k-NN), support vector machine (SVM) and random forest (RF), were employed and also compared their performance based on evaluation criteria.

The k-nearest algorithm is an instance-based learning method used to classify objects based on their closest training examples in the feature space. The k-NN algorithm is implemented using Euclidean distance metrics to locate the nearest neighbor (Boateng *et al.*, 2020). It has some advantages such as having a simple and intuitive algorithm with easy implementation and can be adopted to nonlinear data patterns. In addition, it can be used for both classification and regression (Wang *et al.*, 2024). SVM is another ML method and is used for both classification and regression. It can be applied on both linear and nonlinear data. It employs a margin concept, promoting robustness against overfitting (Wang *et al.*, 2024). The random forest method introduced by Breiman is based on using decision trees in prediction problems. RF, which is an ensemble learning model, obtains the final result by voting or by averaging the results obtained from a large number of decision trees (Breiman, 2001). It is frequently used in the literature due to its advantages. It can be used for both classification or regression and can be applied directly for high-dimensional problems and depends only on one or two tuning parameters. It measures variable importance. Moreover, it is not sensitive to missing values (Cutler *et al.*, 2011).

To determine the contribution of each feature to the model, SHAP that stands for "**SH**apley **A**dditive ex**P**lanations" analysis was conducted. In the RF method, feature importance is given by %IncMSE and IncNodePurity measures but the information regarding the direction (positive/negative) of individual features is not given. On the other hand, SHAP analysis, in addition to information about the importance of the features, provides information about how the features affect the target variable. Also, contribution of the each observation can be determined (Lundberg and Lee, 2017). SHAP gives a more detailed analysis, allowing the capture of nonlinear interactions between target variable and features (Agan *et al.*, 2025).

3.4 Performance evaluation criteria

To evaluate performances of the methods, three evaluation criteria are used, namely root mean squared error (RMSE), mean absolute error (MAE) and R-squared (R^2), respectively. RMSE is calculated as the square root of the MSE and provides a measure of the average magnitude of the prediction error, providing insights into the model's overall accuracy MSE is sensitive to large deviations since it penalizes greater errors more severely (Osman *et al.*, 2025).

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

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

MAE gives a simpler measure of average prediction error, and it is not affected by the magnitude of outliers. A lower RMSE and MAE values indicate that the prediction power of the model gets better.

R^2, the coefficient of determination, gives information about the percentage of variance in the target variable explained by the features. As this value approaches one, it means that the model explains a higher proportion of the variance.

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

where,

y, the actual value of the target variable; $\widehat{y}$, the predicted value of the target variable; $\overline{y}$, the mean value of the target variable and n, the total number of observations.

4. Results

This section gives the results of the analyses for the final data set that consists of 876 country-years observations. Descriptive statistics of the target variable and the features are presented in Table 2.

Table 2. Descriptive statistics

	Mean	Median	Std. deviation	Minimum	Maximum	*n*	VIF
$LNCO_2$	−1.08	−1.24	1.27	−3.51	2.46	876	
LNGDP	7.91	7.80	0.87	6.02	10.34	876	3.3890
LNPop	3.86	3.99	1.25	0.83	6.45	876	1.2057
LNFDI	0.96	0.96	1.11	−2.66	4.64	876	1.2924
LNInflation	1.66	1.74	0.95	−1.77	5.10	876	1.0718
LNTNRR	2.08	2.14	0.95	−2.41	4.03	876	1.8677
REnergy	67.47	77.00	24.66	0.90	98.30	876	1.9719
Urban	40.54	39.49	16.62	8.68	90.42	876	2.1262
Gov	29.16	26.81	17.84	0.31	76.53	876	3.5419
Trade	65.35	62.79	26.78	9.96	171.10	876	1.6394

Source(s): Author's own work

According to the descriptive statistics, the mean of the $LNCO_2$ is found to be a negative value, and its standard deviation is 1.27, while the LNFDI has a mean of 0.96 and a standard deviation of 1.11. Both features have high variability. Regarding other variables, it is seen that LNInflation and Gov have higher variability compared with other features.

After calculating descriptive statistics, Pearson correlation coefficients were calculated between the features. As can be seen from Table 3 CO_2 emissions are related to all features, especially there is positively strong relationship between $LNCO_2$ and LNGDP whereas there is a statistically significant but very week relationship between $LNCO_2$ and LNFDI. Additionally, REnergy and urban features are related with CO_2 strongly. Moreover, there is a similar relationships between LNGDP and REnergy and also Urban.

Table 3. Pearson correlation coefficients for the target variable and the features

	$LNCO_2$	LNGDP	LNPop	LNFDI	LNInf	LNTNRR	REnergy	Urban	Gov	Trade
$LNCO_2$	1	*0.902***	−0.142**	0.102**	−0.132**	−0.206**	*−0.783***	*0.668***	0.460**	0.473**
LNGDP		1	−0.161**	0.097**	−0.136**	−0.207**	*−0.743***	*0.640***	0.437**	0.424**
LNPop			1	−0.172**	0.067*	−0.168**	0.060	−0.272**	0.032	−0.271**
LNFDI				1	0.062	0.252**	−0.123**	0.255**	0.069*	0.331**
LNInf					1	0.117**	0.149**	−0.178**	−0.100**	−0.040
LNTNRR						1	0.327**	0.099**	−0.491**	0.088**
REnergy							1	−0.510**	−0.569**	−0.448**
Urban								1	0.176**	0.368**
Gov									1	0.242**
Trade										1

Note(s): ** and * are significant levels at 1 and 5%, respectively

Source(s): Author's own work

In addition to the correlation coefficients, variance inflation factor (VIF) values were checked to reveal whether there is a multicollinearity problem among features. Since all VIF values are less than 5, it can be said that there is no multicollinearity between the features.

Given the performance evaluation criteria results presented in Table 4, it has been seen that the RF method has the best performance in terms of prediction. After determining the best algorithm, the importance degree of the features was investigated using SHAP analysis.

Table 4. Results of performance evaluation criteria

Method	RMSE	R^2	MAE
k-NN regression	0.1542049	0.9693099	0.1074576
SVM regression	0.1597935	0.9701880	0.1241415
RF regression	0.1213653	0.9834033	0.0919772

Source(s): Author's own work

As can be seen in Figure 3, LNGDP is the most important feature in explaining CO_2 emissions. The second most important feature is REnergy and the third is Urban. While governance, trade openness and LNPop follow these three variables, LNTNRR, LNFDI and LNInflation are the three least important features. It is clear that FDI ranks eighth among the nine features and has no significant impact on CO_2.

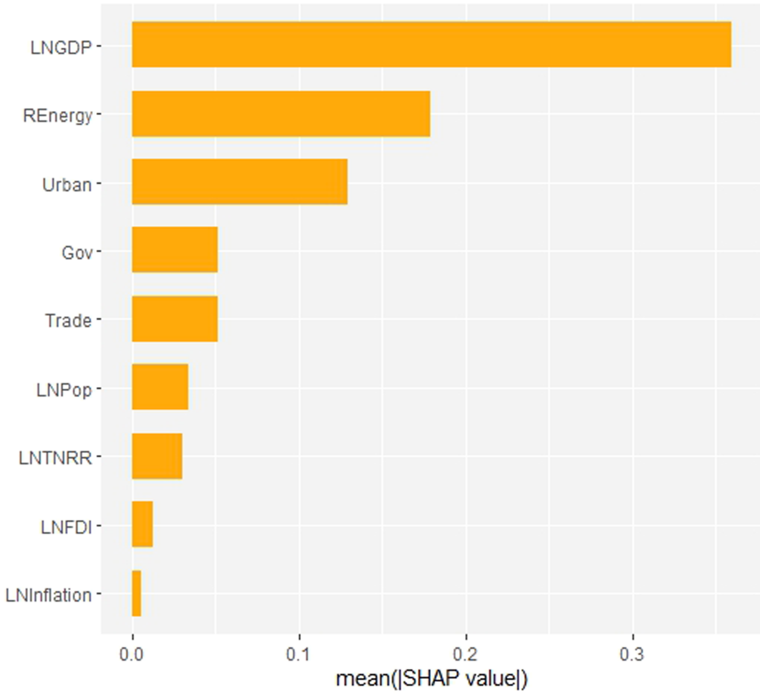

Figure 3. Feature importance in RF method

In Figure 4, the color orange indicates a higher value and purple indicates a lower value. If the feature is related to a reduction in the estimated emissions, a dot on the left side of the figure will be displayed, that is, the feature will have a negative SHAP value. In line with this statement, there is a positive relationship between $LNCO_2$ and LNGDP as well as $LNCO_2$ and Urb. On the other hand, $LNCO_2$ and REnergy are negatively related. According to the same figure, despite the effect is not very high, governance and trade openness are seen to be positively correlated with $LNCO_{2}$..

5. Discussion and conclusion

The global temperature continues to rise, and as a devastating consequences of climate change, the number of disasters such as droughts, water scarcity, severe fires, melting polar ice, catastrophic storms increase throughout the world in the last two decades. To prevent these disasters, about 200 countries adopted the Paris Agreement that aims to limit global temperature increase to 1.5 °C and keep it below 2 °C. However, according to a report published by the UN, 2024 was the hottest year on record, approximately 1.55 °C above pre-industrial levels. This situation indicates that the precautions taken are not sufficient and that environmental degradation is increasing day by day. Depending on this situation, understanding the factors affecting CO_2 emissions plays an important role in preventing environmental degradation.

In this study, whether there is a relationship between FDI and environmental degradation is investigated using ML methods in SSA countries for the period 2002–2021, and possible drivers of environmental degradation are taken into account based on the literature. While CO_2 emissions represents environmental degradation, FDI, GDP per capita, urbanization rate, renewable energy consumption, trade openness, population density, natural resources rents,

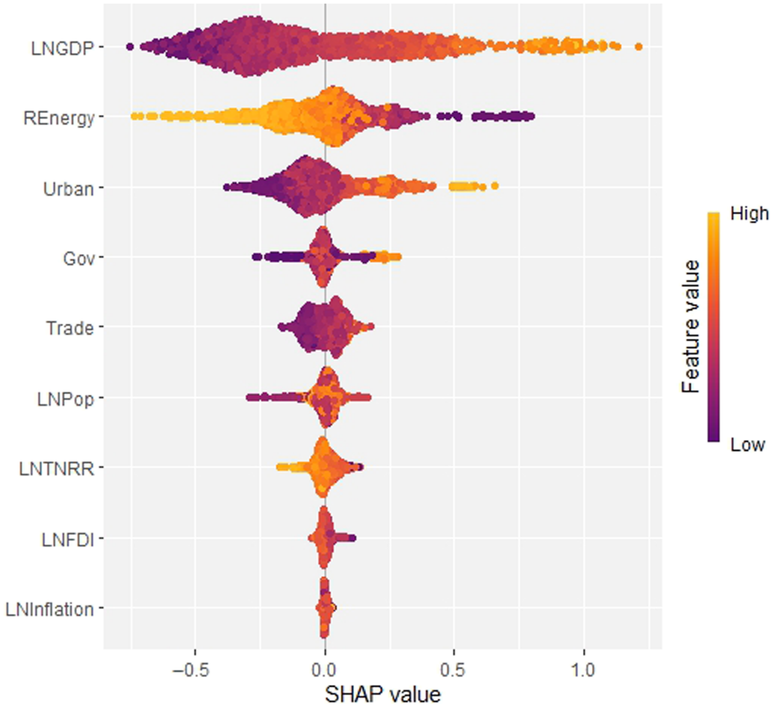

Figure 4. SHAP values for the features

governance and inflation are used as features in the analysis. The presence of multicollinearity among features is examined with VIF values, and it has been seen that there is no multicollinearity between variables. Missing values and outliers, which should be handled at the data preprocessing stage in ML, are taken into account before the analyses and necessary adjustments are made. To determine the relationship between FDI and environmental degradation, econometrics methods have been used in the literature so far. Unlike these studies conducted, k-NN, SVM and RF methods are employed to reveal the impact of the FDI on environmental degradation in this study.

Unlike previous studies, findings obtained indicated that FDI has no significant impact on CO_2 emissions in SSA countries over the analysis period. Similar findings were found for another regions in the literature (Zhang and Zhou, 2016; Solarin and Al-Mulali, 2018; Shao *et al*., 2019; Albulescu *et al*., 2019). With respect to other features, GDP per capita is found to be the most important feature in explaining CO_2 emissions. Renewable energy consumption and urbanization are following this feature, respectively. While GDP per capita and urbanization are positively related with CO_2, there is a negative relationship between renewable energy use and CO_2 emissions.

Numerous studies have found that GDP is the most important variable affecting CO_2 emissions, and this study supports this finding. In contrast, Lawson (2020) has proven that the cause of CO_2 emissions is not economic growth but fossil fuel energy consumption in SSA countries. In addition, other findings of this study support the previous studies that revealed negative effect of the renewable energy consumption on CO_2 and a positive effect of the urbanization on CO_2 (Wang and Dong, 2019; Essandoh *et al*., 2020; Vural, 2020; Bekele *et al*., 2024)). Despite some studies indicated that natural rents are positively linked with environmental degradation in the literature (Sibanda *et al*., 2023; Prempeh *et al*., 2024;

Akadiri *et al*., 2024), the findings of this study show that TNRR has no significant impact on CO_2 but it is clear that these variables are negatively related. With respect to the method used in this study, the findings are consistent with previous studies (Imam *et al*., 2024; Agan *et al*., 2025; Osman *et al*., 2025). Accordingly, RF was found to be the best method in terms of prediction accuracy.

The development of technologies that reduce greenhouse gas emissions, the removal of emissions from the atmosphere and ensuring adaptation to a changing climate are referred to as climate innovation. These activities, which are becoming increasingly important today, are guiding principles for achieving the 2030 SDGs (Inci and Kuang, 2025). From this point, SSA countries can fulfill these activities to ensure environmental sustainability. Additionally, agriculture is one of the most vulnerable areas to climate change, which is vital to the economies of SSA countries. Therefore, developing agricultural policies and implementing fundamental reforms should be essential for SSA countries.

Despite the fact that the amount of emissions is not very high when compared to other countries in the world, CO_2 emissions have increased over time in SSA countries. In addition to this, increasing in environmental degradation contributes to income inequality and poverty. This relationship has been proved in several studies in the literature (Baloch *et al*., 2020; Ssekibaala and Kasule, 2023; Issoufou-Ahmed and Sebri, 2024). Considering SSA countries are mostly low- and lower-middle income countries, it is clear that they will suffer the most from any adverse changes. Therefore, it is vitally important to understand the factors causing environmental degradation in these countries and to take the necessary measures. Accordingly, policymakers should support the use of clean energy technologies and implement regulations to protect the environment. Furthermore, investments in the renewable energy sector should be increased.

While the relationship between FDI and environmental degradation is frequently investigated through econometric analyses in the literature, there are few studies using ML methods. This study fills the gap by using different ML methods to determine this relationship between FDI and environmental degradation. As a further research, other machine learning methods can be used and compared to enhance perspective in this regard or can be used on another group of countries.

Note

1. In this study, 46 SSA countries are included: Angola, Benin, Botswana, Burkina Faso, Burundi, Cabo Verde, Cameroon, Central African Republic, Chad, Comoros, Congo Republic, Congo Democratic Republic, Côte d'ivoire, Equatorial Guinea, Eswatini, Ethiopia, Gabon, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Sudan, Tanzania, The Gambia, Togo, Uganda, Zambia and Zimbabwe.

References

Adams, S. and Klobodu, E.K.M. (2017), "Urbanization, democracy, bureaucratic quality, and environmental degradation", *Journal of Policy Modeling*, Vol. 39 No. 6, pp. 1035-1051, doi: 10.1016/j.jpolmod.2017.04.006.

Agan, B., Celik, S., Damak, O.I. and Miba'am, B. (2025), "Evaluating the machine learning-based models for predicting carbon neutrality in Sub-Saharan African nations", *Environment, Development and Sustainability,* Vol. 27 No. 11, pp. 28185-28219, doi: 10.1007/s10668-025-06289-y.

Akadiri, S.S., Olasehinde-Willams, G., Haouas, I., Lawal, G.O., Fatigun, A.S. and Sadiq Bamgbopa, Y. (2024), Natural resource rent, financial globalization, and environmental degradation: evidence from a resource rich country, *Energy and Environment*, Vol. 35 No. 6, pp. 2911-2934, doi: 10.1177/0958305X231159446

Albulescu, C.T., Tiwari, A.K., Yoon, S.M. and Kang, S.H. (2019), "FDI, income, and environmental pollution in Latin America: replication and extension using panel quantiles regression analysis", *Energy Economics*, Vol. 84, 104504, doi: 10.1016/j.eneco.2019.104504.

Alhassan, H. and Kwakwa, P.A. (2023), "The effect of natural resources extraction and public debt on environmental sustainability", *Management of Environmental Quality: An International Journal*, Vol. 34 No. 3, pp. 605-623, doi: 10.1108/meq-07-2022-0192.

Alpaydın, E. (2020), *Introduction to Machine Learning*, 4th ed., The MIT Press, Cambridge, Massachusetts London.

Asrol, M., Papilo, P. and Gunawan, F.E. (2021), "Support vector machine with K-fold validation to improve the industry's sustainability performance classification", *Procedia Computer Science*, Vol. 179, pp. 854-862, doi: 10.1016/j.procs.2021.01.074.

Baloch, M.A.D., Khan, S.U., Ulucak, Z.Ş. and Ahmad, A. (2020), "Analyzing the relationship between poverty, income inequality, and CO_2 emission in Sub-Saharan African countries", *Science of the Total Environment*, Vol. 740, 139867, doi: 10.1016/j.scitotenv.2020.139867.

Bekele, M., Sassi, M., Jemal, K. and Ahmed, B. (2024), "The dynamic linkage between renewable energy consumption and environmental sustainability in Sub-Saharan African countries: heterogeneous macro-panel data analysis", *Cogent Economics and Finance*, Vol. 12 No. 1, 2285188, doi: 10.1080/23322039.2023.2285188.

Boateng, E.Y., Otoo, J. and Abaye, D.A. (2020), "Basic tenets of classification algorithms K-nearest-neighbor, support vector machine, random forest and neural network: a review", *Journal of Data Analysis and Information Processing*, Vol. 8 No. 04, pp. 341-357, doi: 10.4236/jdaip.2020.84020.

Bokpin, G.A. (2017), "Foreign direct investment and environmental sustainability in Africa: the role of institutions and governance", *Research in International Business and Finance*, Vol. 39, pp. 239-247, doi: 10.1016/j.ribaf.2016.07.038.

Breiman, L. (2001), "Random forests", *Machine Learning*, Vol. 45 No. 1, pp. 5-32, doi: 10.1023/a:1010933404324.

Burck, J., Uhlich, T., Bals, C., Höhne, N., Nascimento, L., Kumar, C.H., Bosse, J., Riebandt, M. and Pradipta, G. (2023), "Climate change performance index 2024, December, Germanwatch", New Climate Institute and Climate Action Network.

Chien, F.S., Chau, K.Y. and Sadiq, M. (2023), "Impact of climate mitigation technology and natural resource management on climate change in China", *Resources Policy*, Vol. 81, 103367, doi: 10.1016/j.resourpol.2023.103367.

Cutler, A., Cutler, R.D. and Stevens, J.R. (2011), "Random forests", *Chapter in Machine Learning*, pp. 157-175, doi: 10.1007/978-1-4419-9326-7_5.

Dhrifi, A., Jaziri, R. and Alnahdi, S. (2020), "Does foreign direct investment and environmental degradation matter for poverty? Evidence from developing countries", *Structural Change and Economic Dynamics*, Vol. 52, pp. 13-21, doi: 10.1016/j.strueco.2019.09.008.

Dua, W.C. and Xiab, X.H. (2018), "How does urbanization affect GHG emissions? A cross-country panel threshold data analysis", *Applied Energy*, Vol. 229, pp. 872-883, doi: 10.1016/j.apenergy.2018.08.050.

Duan, Y. and Jiang, X. (2021), "Pollution haven or pollution halo? A re-evaluation on the role of multinational enterprises in global CO_2 emissions", *Energy Economics*, Vol. 97, 105181, doi: 10.1016/j.eneco.2021.105181.

Espoir, D.K. and Oyadeyi, O.O. (2025), "Investigating the pollution haven hypothesis in the top-5 FDI recipient countries in Africa: evidence from heterogeneous panel data with sharp and smooth transitions", *Journal of Environmental Management*, Vol. 392, 126536, doi: 10.1016/j.jenvman.2025.126536.

Espoir, D.K., Mudiangombe, B.M., Bannor, F., Sunge, R. and Tshitaka, J.M. (2022), "CO_2 emissions and economic growth: assessing the heterogeneous effects across climate regimes in Africa", *Science of the Total Environment*, Vol. 804, 150089, doi: 10.1016/j.scitotenv.2021.150089.

Espoir, D.K., Sunge, R. and Bannor, F. (2023), "Economic growth, renewable and nonrenewable electricity consumption: fresh evidence from a panel sample of African countries", *Energy Nexus*, Vol. 9, 100165, doi: 10.1016/j.nexus.2022.100165.

Espoir, D.K., Sunge, R., Nchofoung, T. and Alola, A.A. (2024), "Analysing the drivers of ecological footprint in Africa with machine learning algorithm", *Environmental Impact Assessment Review*, Vol. 10, 107332, doi: 10.1016/j.eiar.2023.107332.

Essandoh, O.K., Islam, M. and Kakinaka, M. (2020), "Linking international trade and foreign direct investment to CO2 emissions: any differences between developed and developing countries?", *Science of the Total Environment*, Vol. 712, 136437, doi: 10.1016/j.scitotenv.2019.136437.

Güney, T. (2017), "Governance and sustainable development: how effective is governance?", *Journal of International Trade and Economic Development*, Vol. 26 No. 3, pp. 316-335, doi: 10.1080/09638199.2016.1249391.

Güney, T. (2019), "Renewable energy, non-renewable energy and sustainable development", *The International Journal of Sustainable Development and World Ecology*, Vol. 26 No. 5, pp. 389-397, doi: 10.1080/13504509.2019.1595214.

Güney, T. (2022), "Solar energy, governance and CO_2 emissions", *Renewable Energy*, Vol. 184, pp. 791-798, doi: 10.1016/j.renene.2021.11.124.

Huang, X., Ahmad, W., Umair, M., Antohi, V.M., Fortea, C. and Cristache, N. (2025), "Foreign direct investment, trade openness and environmental pollution in Pakistan: does renewable energy mitigate environmental degradation?", *Frontiers Enviromental Science*, Vol. 13, 1618767, doi: 10.3389/fenvs.2025.1618767.

Imam, M., Adam, S., Dev, S. and Nesa, N. (2024), "Air quality monitoring using statistical learning models for sustainable environment", *Intelligent Systems with Applications*, Vol. 22, 200333, doi: 10.1016/j.iswa.2024.200333.

Inci, A.C. and Kuang, H. (2025), "Sustainable finance for sustainable development regulations", in Aras, G. (Ed.), *Theory and Practice*, Routledge, pp. 58-85.

Issoufou-Ahmed, O. and Sebri, M. (2024), "On the transmission channels driving climate change-income inequality nexus in sub-Saharan African countries", *Business Strategy and Development*, Vol. 7 No. 1, e339, doi: 10.1002/bsd2.339.

Karim, S., Appiah, M., Naeem, M.A., Brian, M.L. and Li, M. (2022), "Modelling the role of institutional quality on carbon emissions in Sub-Saharan African countries", *Renewable Energy*, Vol. 198, pp. 213-221, doi: 10.1016/j.renene.2022.08.074.

Khan, K.A., Anser, M.K., Pala, F., Barut, A. and Zafar, M.W. (2024), "Promoting sustainability in developing countries: a Machine Learning-based approach to understanding the relationship between green investment and environmental degradation", *Gondwana Research*, Vol. 132, pp. 136-149, doi: 10.1016/j.gr.2024.03.013.

Koirala, B.S. and Pradhan, G. (2020), "Determinants of sustainable development: evidence from 12 Asian countries", *Sustainable Development*, Vol. 28 No. 1, pp. 39-45, doi: 10.1002/sd.1963.

Kutlu, F.O. and Atis, S. (2024), "Does foreign direct investment affect environmental degradation: evidence from largest carbon intense countries", *PLoS One*, Vol. 19 No. 11, e0314232, doi: 10.1371/journal.pone.0314232.

Lawson, L.A. (2020), "GHG emissions and fossil energy use as consequences of efforts of improving human well-being in Africa", *Journal of Environmental Management*, Vol. 273, 111136, doi: 10.1016/j.jenvman.2020.111136.

Li, Y., Zhou, R., Zhou, Y. and Chen, Z.S. (2025), "Machine learning techniques in maritime environmental sustainability: a comprehensive review of the state of the art", *Computers and Electrical Engineering*, Vol. 124, 110395, doi: 10.1016/j.compeleceng.2025.110395.

Lundberg, S.M. and Lee, S. (2017), "A unified approach to interpreting model predictions", *31st Conference on Neural Information Processing Systems*, Long Beach, CA.

Osman, B.M., Mohamud, H., Omar, A.F., Yabarow, A.A., Omar, O.M., Jirow, M.S.A., Dahir, A.N. and Nur, A.M. (2025), "The impact of climate change on economic growth in Somalia: using random forest and Bayesian approach", *Cogent Economics and Finance*, Vol. 13 No. 1, 2496684, doi: 10.1080/23322039.2025.2496684.

Ploeg, F. (2011), "Natural resources: curse or blessing?", *Journal of Economic Literature*, Vol. 49 No. 2, pp. 366-420, doi: 10.1257/jel.49.2.366.

Prempeh, K.B., Kyeremeh, C., Yeboah, S.A. and Danso, F.K. (2024), "Quantifying carbon emissions through financial development in Ghana: empirical evidence from novel dynamic ARDL and KRLS techniques", *Cogent Economics and Finance*, Vol. 12 No. 1, 2423260, doi: 10.1080/23322039.2024.2423260.

Safdar, S., Khan, A. and Andlib, Z. (2022), "Impact of good governance and natural resource rent on economic and environmental sustainability: an empirical analysis for South Asian economies", *Environmental Science and Pollution Research*, Vol. 29 No. 55, pp. 82948-82965, doi: 10.1007/s11356-022-21401-9.

Sakariyahu, R., Fagbemi, T., Adigun, R., Lawal, R., Seyingbo, O. and Oyekola, O. (2024), "Severity of environmental degradation and the impact on quality of life in Africa", *Journal of Environmental Management*, Vol. 356, 120537, doi: 10.1016/j.jenvman.2024.120537.

Sánchez, M.G. and Ortega, J.L.M. (2020), "Greenhouse gas emissions growth in Europe: a comparative analysis of determinants", *Sustainability*, Vol. 12 No. 3, p. 1012, doi: 10.3390/su12031012.

Shao, Q., Wang, X., Zhou, Q. and Balogh, L. (2019), "Pollution haven hypothesis revisited: a comparison of the BRICS and MINT countries based on VECM approach", *Journal of Cleaner Production*, Vol. 227, pp. 724-738, doi: 10.1016/j.jclepro.2019.04.206.

Sibanda, K., Garidzirai, R., Mushonga, F. and Gonese, D. (2023), "Natural resource rents, institutional quality, and environmental degradation in resource-rich sub-Saharan African countries", *Sustainability*, Vol. 15 No. 2, p. 1141, doi: 10.3390/su15021141.

Solarin, S.A. and Al-Mulali, U. (2018), "Influence of foreign direct investment on indicators of environmental degradation", *Environmental Science and Pollution Research*, Vol. 25, pp. 24845-24859, doi: 10.1007/s11356-018-2562-5.

Ssekibaala, S.D. and Kasule, T.A. (2023), "Examination of the poverty-environmental degradation nexus in Sub-Saharan Africa", *Regional Sustainability*, Vol. 4 No. 3, pp. 296-308, doi: 10.1016/j.regsus.2023.08.007.

Statista (2024), "Politics and society, global climate change", available at: https://www-statista-com/10974/climate-change–statista-dossier

Statista (2025), "Climate change in Africa: a statista dossıer plus on the increasıng threat of climate change to Afrıca's development".

Unites Nations (2025), "The sustainable development goals report 2025: special edition towards a rescue plan for people and planet".

Vural, G. (2020), "How do output, trade, renewable energy and non-renewable energy impact carbon emissions in selected Sub-Saharan African Countries?", *Resources Policy*, Vol. 69, 101840, doi: 10.1016/j.resourpol.2020.101840.

Wang, J. and Dong, K. (2019), "What drives environmental degradation? Evidence from 14 Sub-Saharan African countries", *Science of the Total Environment*, Vol. 656, pp. 165-173, doi: 10.1016/j.scitotenv.2018.11.354.

Wang, W., Dezhi, L., Zhou, S., Wang, Y. and Yu, L. (2024), "Exploring the key influencing factors of low-carbon innovation from urban characteristics in China using interpretable machine learning", *Environmental Impact Assessment Review*, Vol. 107, 107573, doi: 10.1016/j.eiar.2024.107573.

World Bank (2025a), "World development indicators database", available at: https://databank.worldbank.org/source/world-development-indicators (accessed July 2025).

World Bank (2025b), "World governance indicators database", available at: https://databank.worldbank.org/source/worldwide-governance-indicators (accessed July 2025).

Zhang, C. and Zhou, X. (2016), "Does foreign direct investment lead to lower CO2 emissions? Evidence from a regional analysis in China", *Renewable and Sustainable Energy Reviews*, Vol. 58, pp. 943-951, doi: 10.1016/j.rser.2015.12.226.

Corresponding author
Nuray Tezcan can be contacted at: nuraytezcan@halic.edu.tr