The Industry of the Future: From Industry 4.0 to Industry 5.0 – Integration of Humans and Technology: New Technologies

Karabegović, Isak

2025
Akademija nauka i umjetnosti Bosne i Hercegovine
https://bastina.anubih.ba/handle/123456789/837

DOI: 10.5644/PI2025.220.12

Machine Learning Modelling of CO_2 Emissions for Sustainable Development: A Comparative Study of Bosnia and Herzegovina, Croatia, and Slovenia

Mirza Pašić[*1], Nejra Biber[1,2]

Abstract: *The growing need for reducing CO_2emissions in the context of sustainable development has intensified the search for efficient analytical approaches to understand and manage emission drivers. In this paper, three machine learning models were developed using multiple linear regression for the countries of Bosnia and Herzegovina, Croatia and Slovenia. Renewable energy consumption, $PM_{2,5}$ air pollution, GDP per capita, foreign direct investment, urban population, forest area, and total population were used as inputs in the models, while CO_2 emissions for the period from 2000 to 2020 were used as outputs. The developed models for all three countries have good performance, with R^2values of 91,34%, 77,91%, and 77,20% respectively. For Bosnia and Herzegovina urban population increases CO_2emission, while renewable energy consumption and forest area decrease CO_2 emission. In Croatia $PM_{2,5}$ was the most influential factor that increases CO_2emission.In Slovenia population growth decreases CO_2 emissions, whileGDP per capita increases CO_2 emissions. Also, hypothesis testing for differences between means was performed for all variables between all three countries. The findings showed that for almost all variables there were statistically significant differences in mean differences between all countries. Regarding CO_2 emission there are not enough statistical evidence that Bosnia and Herzegovina have higher CO_2 emissions than Croatia, while both Bosnia and Herzegovina, and Croatia have significantly higher CO_2 emissions than Slovenia. This research shows the potential of machine learning models as tools for data-driven policymaking in the transition towards Industry 5.0 and a sustainable industrial future.*

Keywords: *CO_2 emission, sustainability, machine learning, multiple linear regression, hypothesis tests*

1. Introduction

The increase in greenhouse gas (GHG) emissions, primarily caused by human activities, is responsible for global warming and climate change, which is why the urgent action is needed to avoid potentially irreversible damage to the environment [1]. Climate change is a global problem whose consequences

[*1]University of Sarajevo,Faculty of Mechanical Engineering, Sarajevo, Bosnia and Herzegovina
[2]Center for Economic, Technological and Environmental Development Sarajevo - CETEOR, Bosnia and Herzegovina
E-mail: mirza.pasic@mef.unsa.ba

negatively affect humanity, biodiversity, terrestrial and marine ecosystems, but also peace and security [2]. The effects of climate change call for strengthening global responses with mitigation and adaptation actions to keep a global temperature rise below 2,0 degrees Celsius above pre-industrial levels and to pursue efforts to limit the temperature increase even further to 1,5 degrees Celsius, as set out in the Paris Agreement goals [2, 3]. The number of countries, cities, businesses and other institutions pledging to achieve net zero emissions grows constantly. However, the transition to a net-zero world is considered to be one of the greatest challenges as it requires the complete transformation of how the population produce, consume, and move about [4].

Under the EU Climate Law, the European Green Deal aims to make Europe the first climate-neutral continent by 2050, starting with the aim to reduce GHG emissions by at least 55% by 2030 compared to 1990 levels [5].EU aims to achieve climate goals and encourage transition in both EU and non-EU countries by introducing different measures through its legislation and mechanisms such as EU Emissions Trading System (ETS) and Carbon Border Adjustment Mechanism (CBAM). 'Fit for 55' package of legislation introduces reforms across all sectors of the EU's economy with measures for promoting zero-emission mobility, developing a net-zero industry, building a clean energy system based on renewable energy sources (RES), renovating buildings for energy efficiency, restoring nature and enabling biodiversity, all to accelerate the transition towards climate neutrality [5]. An important instrument for achieving EU climate goals in a cost-effective way is the EU ETS. As the world's first carbon market that currently operates in all EU countries, Iceland, Liechtenstein and Norway, and is linked to the Swiss ETS, EU ETS collects revenue that must be used to support the green transition by investing in renewable energy, energy efficiency, improvements and low-carbon technologies that help reduce emissions [6]. From 2026, EU will also apply the CBAM to prevent the risk of "carbon leakage" from countries with less stringent climate policies than in the EU. With CBAM, EU aims to put a fair price on the GHG emissions emitted during the production of carbon intensive goods and ensure that the carbon price of imports is equivalent to the carbon price of domestic production [7].

Total GHG emissions are the sum of emissions of various gases, but the most dominant one is carbon dioxide (CO_2). Total CO_2 emissions accounted for 74,89% of total GHG emissions in 2023 and amounted to 41,42 billion tons in the same year [8, 9].In advanced economiesCO_2emissions decrease, while in most emergingmarkets and developing economies CO2 emissions continue to grow, but more slowly due to a rise in clean energy deployment[10].

The divergence in emissions trends between different regions has led many organizations and researchers to explore the factors influencing $CO_2$and overall GHG emissions and their impact on sustainable development.Authors in [11] focused on analyzing the effects of GDP, renewable energy, household energy consumption and waste on the GHG emissions and found a positive relationship between real GDP, household energy consumption and waste generation and GHG emissions, while the share of renewable energy turned out to be slightly negative.The impact of artificial intelligence (AI), economic growth, FDI, energy consumption, and urbanization on CO_2emissions was investigated in [12]. The analysis showed that AI positively correlates with CO_2reduction, while GDP growth, energy consumption, FDI, and urbanization intensify CO_2emissions. Authors in [13] investigated the relationship between economic growth and CO_2 emissions worldwide using a multilevel model that accommodates interactions between fixed and random effects related to GDP and CO_2emissions. The results of the analysis showed a positive relationship between economic development and CO_2 emissions.

Because the precise forecasting plays an important role in regulating reductions in CO_2 emissions, the authors in [14] developed four machine learning algorithms to estimate CO_2 emissions based on the data on different energy sources that impact CO_2 emissions.In [15] authors did bibliographic analysis and content review to analyze the how using machine learning can be used in order to reduce CO_2emissions in construction. The research showed that different machine learning model were used such as artificial neural networks, genetic algorithms, regression models, support vector machines, and decision trees, were used to predict CO_2 emissions and to optimize sustainable construction.Study [16] used machine learning methods gradient boosting machine, support vector machines, random forest, extreme gradient boosting to predict city-level CO_2 emissions throughout China. The authors concluded that their findings can be used to create carbon footprint maps.The authors in [17] used AI and machine learning to predict CO_2 emissions in the near and far future, and different models were created to analyze CO_2 emission during and after COVID-19 pandemic.

In [18] the effects of economic growth, industrial structure, urbanization, investments in research and development, use of foreign capital, and energy consumption on CO_2emissions using machine learning algorithms was analyzed. The results showed the importance of region-specific industrialization and an optimal urbanization range as well as an increased research and development and foreign capital investment in green technologies to support the reductions in CO_2emissions, while sustaining economic growth. Magazzino et al. [19] did a study in which authors did a machine learning approach to analyze solar and wind energy production, coal consumption, GDP, and CO_2emissions. This research was done to analyzeCO_2 emissions for China, India, and USA and it was concluded that India has a potential for increased CO_2 emissions. USA and China demonstrate a stronger potential for achieving sustainability than India.Authors used machine learning methods such as adaptive boosted support vector machines (SVR) and Adaptive boosted Gaussian process (GPR) to predict loading of CO_2in the solvent phase. In this research temperature, solution concentration, CO_2 particle pressure, amino acid salt molecular weight, melting point of amino acid salt, molecular weight of cation were used as inputs, while loading of CO_2 in the solvent phase was used as output [20].Linear regression approach was used in [21] to explore the relationship between CO_2 emissions, energy use, GDP, and population. The findings of this case study showed a long-run equilibrium relationship from these factors to CO_2 emissions, and that an increase of 1% in energy use, GDP and population leads to an increase of 0,58%, 0,73%, and 1,30% in CO_2emissions, respectively. A comparison of conventional linear regression model and linear regression model with fuzzy numbers was done in [22] to predict CO_2 emissions, using data on CO_2emissions, fuel mix, transportation, GDP, and population. The results showed that multiple linear regression performed better in predicting CO_2emissions.

Previously analysed research show that machine learning and artificial intelligence methods are effective tools for predicting CO_2 emissions, enabling a better understanding of the factors contributing to emissions and the identification of strategies to reduce them. Technologies such as automation, the Internet of Things (IoT), big data, AI and machine learning, associated with Industry 4,0, allowed industries to achieve higher levels of efficiency and productivity. However, even though the embracement of Industry 4,0 led to considerable improvements in performances, the industries globally are embracing the concepts of Industry 5.0 which emphasizes collaboration between humans and machines to achieve more flexible, adaptable, and sustainable systems. Industry 5.0 focuses more on sustainable practices to optimize resource use and lower environmental impacts, and on achieving more inclusive and responsible industrial practices [23]. This is in line with the global goals to reduce emissions and promote environmental sustainability.

In [24] authors identified 16 interdependent functions through which Industry 5.0 contributes to sustainability. Authors also emphasized that experts suggested that Industry 5.0 has been developed as a complement to Industry 4,0 to address social and environmental challenges brought by the digital transformation. Energy efficiency and renewable energy in Industry 5.0 and digital industrial revolution from the perspective of sustainability was analyzed in [25], while the need for mass personalization for sustainability and Industry 5.0 builds upon Industry 4,0 by introducing a human-centric approach that promotes greater resilience and sustainability [26]. In [27] principal component regression machine learning model was used to predict levels of heavy metals like lead, zinc, nickel, arsenic, and cadmium, analyze the environmental data, and dealing with the variability in environmental data. Authors suggest that machine learning methods and decision analytics should be incorporated with resilient and human-centered approach in order to be aligned with goals of Industry 5.0. Research [28] aimed to address major challenges of the integration of human-centered AI methodology relevant to circular economy and Industry 5.0.

Various machine learning methods were used forprediction and analysis ofCO_2 emissions, and that in the context of Industry 5.0 where sustainability and the application of advanced technologies such as machine learning are key principles, the importance of this approach should be further emphasized. The aim of this research was to develop machine learning linear regression models for the countries of Bosnia and Herzegovina, Croatia, and Slovenia, as well as to analyze the impact of different statistically significant variables for each country individually. Also, hypothesis tests were performed to analyze the differences between means of all variables used in this research between each country. The structure of the paper is as follows: after the introduction, materials and methods are presented in Section 2. The results and discussion are given in Section 3. Finally, Section 4 presents the conclusion of the research with recommendations for future research.

2. Materials and Methods

In this research average yearly values from 2000 to 2020 were used for the countries of Bosnia and Herzegovina (BIH), Croatia (CRO) and Slovenia (SVN). Multiple linear regression models with backward analysis were developed for all three countries where:

- renewable energy consumption (% of total final energy consumption) – REC,
- $PM_{2,5}$ air pollution (mean annual exposure $\left[\frac{\mu g}{m^3}\right]$) –$PM_{2,5}$,
- GDP per capita (current US$) – $GDPpc$,
- foreign direct investment, net inflows (BoP, current US$) – FDI,
- urban population (% of total population) – UP,
- forest area [km^2]– FA, and
- total population – TP,

were used as independent variables, while CO_2 emissions (total) excluding LULUCF (Mt CO2e) –CO_2 emissionswas used as dependent variable.Also, hypothesis tests for the differences between two means for all variables between BIH, CRO and SVN were done.

2.1. Data

The data for this research were collected from World Bank [29]. Data for all variables were shown in Figure 1, Figure 2 and Figure3 for BiH, Croatia and Slovenia respectively.

From Figure 1 it can be seen that renewable energy consumption were increasing in years 2018, 2019 and 2020, while the values of $PM_{2,5}$were decreasing throughout the years. GDP per capita and urban population were increasing, while the total population was decreasing for Bosnia and Herzegovina.

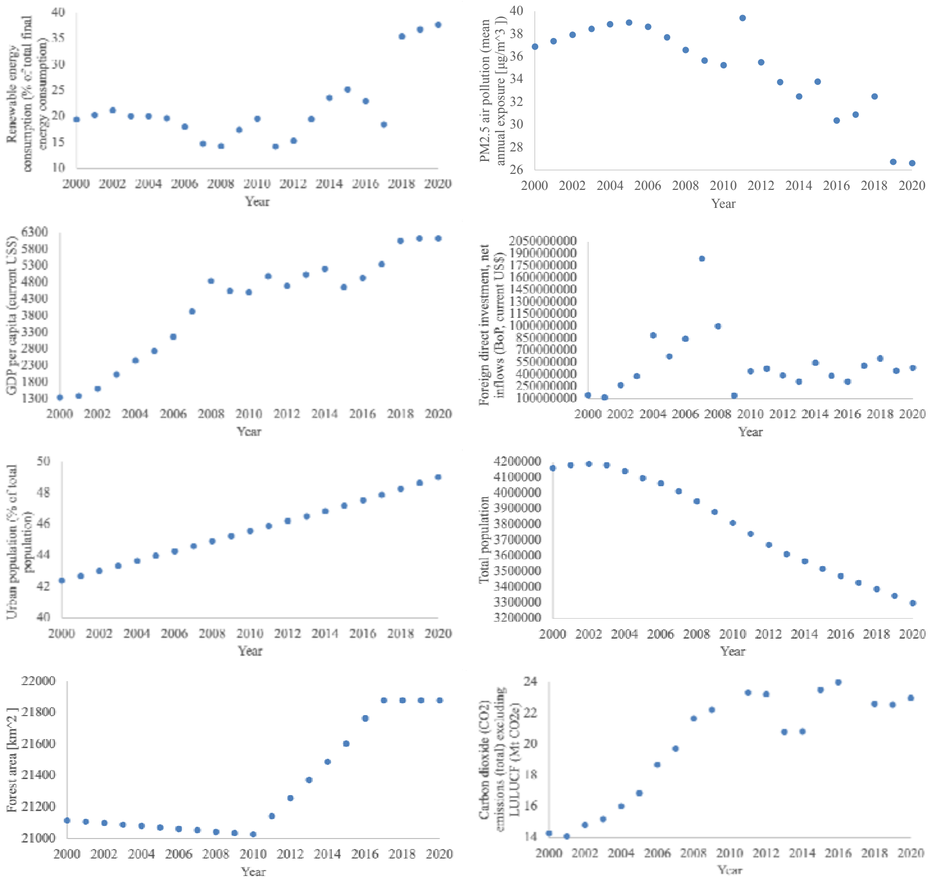

Figure 1. Scatter plots of all variables for Bosnia and Herzegovina for the period from 2000 to 2020

From Figure 2 it can be seen that the values of renewable energy consumption, GDP per capita, urban population and forest were increasing, while for Croatia the values of $PM_{2,5}$, total population, and CO_2 emissions were decreasing over the time period of 2000 – 2020.

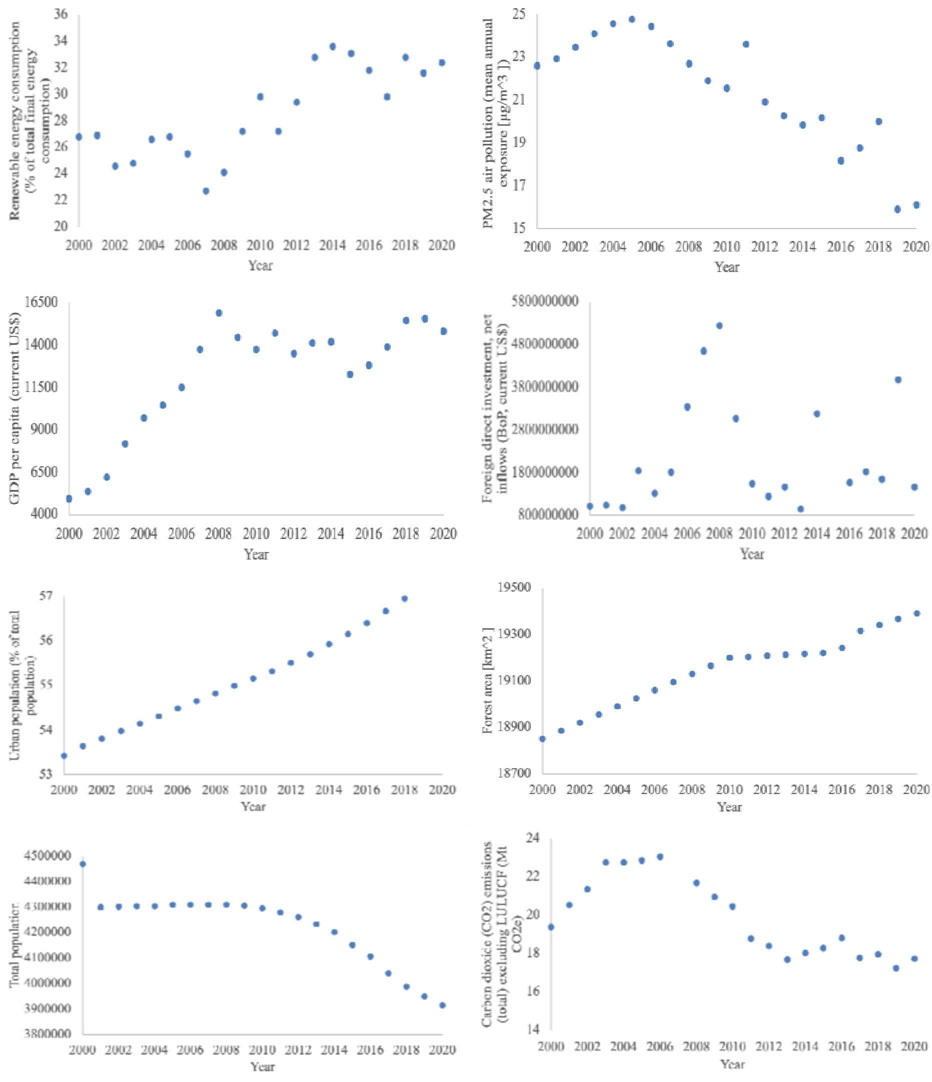

Figure 2. Scatter plots of all variables for Croatia for the period from 2000 to 2020

Figure 3 shows that for Slovenia renewable energy consumption, GDP per capita, urban population, and total population were increasing, while$PM_{2,5}$ and

CO_2emissions were decreasing through the years. Forest area was increasing until 2016 when it started decreasing until 2020.

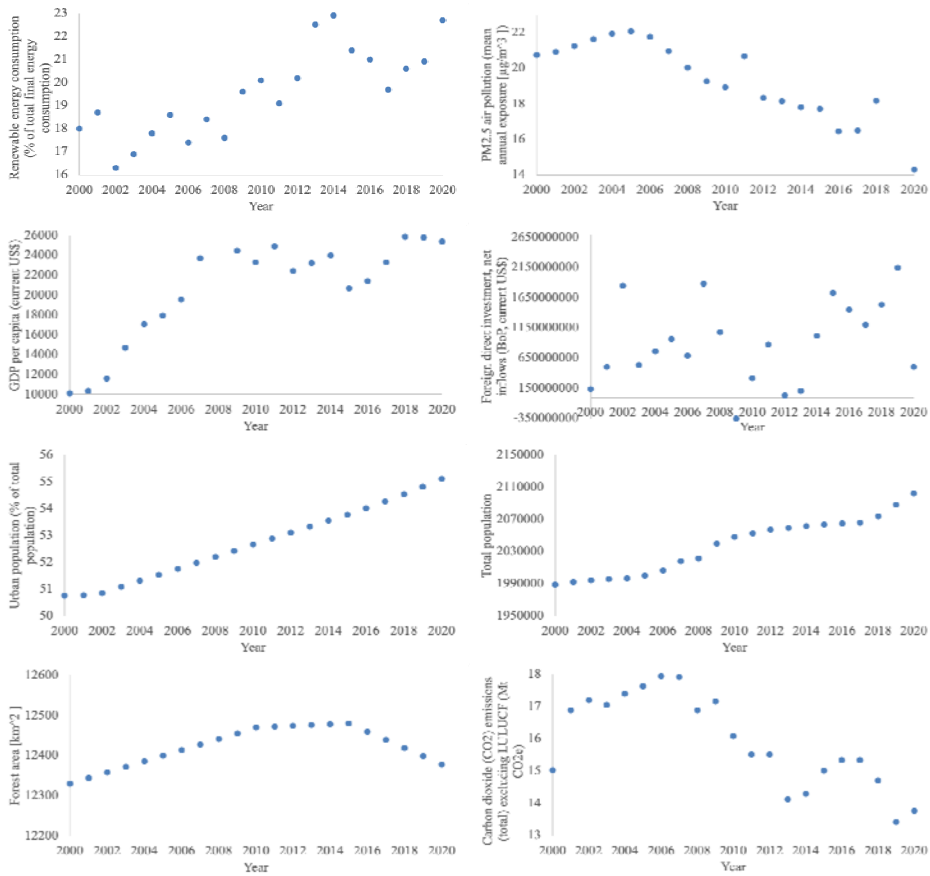

Figure 3. Scatter plots of all variables for Slovenia for the period from 2000 to 2020

Based on the Figure 1, Figure 2, and Figure 3 it is evident that the same variables show different trends for Bosnia and Herzegovina, Croatia and Slovenia, with some variables showing increasing while others showing a decreasing pattern over the years.

2.2.Multiple Linear Regression

In this research multiple linear regression with backward elimination was done. Due to the large differences in the order of magnitude among the variables, data were standardized to reduce potential numerical instability and to make the coefficients of linear regression easier to interpret. Data were analyzed using multiple linear regression modelsshown in equation (1) with the level of significance α= 0,05.

$$\hat{y}_j = b_0 + \sum_{i=1}^{k} b_i x_{ij} \tag{1}$$

where:
k − number of independent variables,
$\hat{y}_j$ − predicted value of dependent variable ($j = 1, \dots, n$),
n − number of samples,
b_0 −intercept,
b_i − regression coefficients ($i = 1, \dots, k$).

Significance tests were carried out for each stage of the backward regression analysis in order to provide statistical evidence regarding the independent variables that significantly contribute to the prediction of theCO_2 emissions. Variance Inflation Factor (VIF) was used to determine potential presence of multicollinearity between all significant independent variables. This calculation of VIF is shown in equation (2), where VIF value lower than 10 indicates that there is no multicollinearity between the significant independent variables[30].

$$VIF = \frac{1}{1 - R_i^2} \tag{2}$$

where:
R_i^2 −coefficient of determination when the i^{th} independent variable is regressed against all other independent variables.
In order to identify the most impactful independent variable,standardized regression coefficients can be calculated as shown in equation (3).

$$b_{ist} = \frac{s_{x_i}}{s_{x_i}} \cdot b_i \tag{3}$$

where:
i − i^{th} independent variable (i = 1, …k),
b_{ist} − standardized regression coefficient,
s_{x_i} − standard deviation of i^{th} independent variable,
s_y − standard deviations of dependent variable.

As standardized values of all variables were used in this study, calculated regression coefficients are standardized regression coefficients.

For each country, multiple linear regression model using backward elimination was evaluated using coefficient of determination(R^2)and mean squared error (MSE) as shown in equation (4) and equation (5) respectively.

$$R^2 = 1 - \frac{\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2}{\sum_{j=1}^{n}\left(y_j - \bar{y}_j\right)^2} \tag{4}$$

where:
$\hat{y}_j -$ predicted value of dependent variable ($j = 1, \dots, n$),
y_j – actual value of dependent variable,
$n -$ number of samples.

$$MSE = \frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2 \tag{5}$$

where:
$\hat{y}_j -$ predicted value of dependent variable ($j = 1, \dots, n$),
y_j – actual value of dependent variable,
$n -$ number of samples.

2.2.Hypothesis tests for the differences between two means for all variables between Bosnia and Herzegovina, Croatia and Slovenia

Hypothesis tests for the differences between two means for all variables were done in order to analyze whether there were differences between the means of the same variable for three countries, BIH, CRO, and SVN. Since in this research yearly average values have been used and the analyzed period is from 2000 to 2020, there are 20 samples per variable for each country. Since population standard deviations are unknown, the $t - test$was applied to test the difference between means. To determine whether to use $t\text{–} test$ that assumes equal variances (pooled variance $t - test$) or the one that does not, $F - test$was performed to assess whether the variances between two samples are different. The F statistic (F_{stat}) is calculated as shown in equation (6):

$$F_{stat} = \frac{s_1^2}{s_2^2} \tag{6}$$

where s_1^2 and s_2^2 are the sample variances of the two groups (with s_1^2>s_2^2).

If the $F - test$shows that the variances are not significantly different, the pooled variance $t - test$ is applied. The t-statistic (t_{stat}) for the pooled variance version is calculated by equation (7):

$$t_{stat} = \frac{\overline{x}_1 - \overline{x}_2}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{7}$$

where s_p is pooled variance calculated by equation (8):

$$s_p = \sqrt{\frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}} \tag{8}$$

If the $F - test$showed unequal variances, then the $t_{stat}$was calculated as shown in equation (9):

$$t_{stat} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^1}{n_1} + \frac{s_2^2}{n_2}}} \tag{9}$$

3. Results and Discussion

In this research multiple linear regression models with backward elimination were created for all three countries, Bosnia and Herzegovina, Croatia, and Slovenia. Renewable energy consumption (% of total final energy consumption) (REC), $PM_{2,5}$ air pollution (mean annual exposure $\left[\frac{\mu g}{m^3}\right]$) ($PM_{2,5}$), GDP per capita (current US$) ($GDPpc$), foreign direct investment, net inflows (BoP, current US$) ($FDI$),urban population (% of total population) (UP), forest area [km^2] (FA), and total population (TP), were used as independent variables, whileCO_2 emissions (total) excluding LULUCF (Mt CO2e)(CO_2) was used as dependent variable. The aim of creating these models was to analyse which independent variables are statistically significant in the models and to analyse the impact of statistically significant variables on models' output CO_2 emissions for each country individually.

Also, hypothesis tests for the differences between two means for all variables between BIH, CRO and SVN were done. Since in this research there were 20 samples for each variable, first $F - test$ was performed to analyze the difference between variances for each variable, and after that $t - test$ was done.

3.1.Analysis of developed linear regression models with backward elimination for each country

Developed linear regression models for Bosnia and Herzegovina, Croatia, and Slovenia are shown in equations (10 – 12) respectively. Since the variables were of different orders of magnitude, their values were standardized prior to developing the linear regression model and thus standardized regression coefficients were obtained. Standardized coefficients indicate how many standard deviations the dependent variable changes when an independent variable changes by one standard deviation. Independent variables with the largest standardized coefficients in absolute terms have the greatest influence on the dependent variable.

1.1 $$CO_{2\,BIH} = -0{,}317 \cdot REC_{BIH} + 1{,}514 \cdot UP_{BIH} - 0{,}551 \cdot FA_{BIH} \quad (10)$$

$$CO_{2\,CRO} = 0{,}813 \cdot PM_{2{,}5\,CRO} + 0{,}355 \cdot FDI_{CRO} \quad (11)$$

$$CO_{2\,,5\,SVN} = 0{,}573 \cdot GDPpc_{SVN} - 1{,}240 \cdot PT_{SLO} \quad (12)$$

From equations (10 – 12) it can be seen that for all three countries different independent variables are statistically significant. These models demonstrate that the analysis of carbon emissions in the context of transitioning from Industry 4,0 to Industry 5.0 must be considered individually for each country, as the context is highly complex. Equation (10) for BIH shows that UP has the greatest influence on CO_2 emissions, indicating that as UP increases, CO_2 emissions also increase. Increase in values of REC and FA lead to decrease in CO_2 emissions, showing that BIH should focus more on renewable energy consumption and increase of forest area in order to achieve decarbonization and CO_2emission reduction.In CRO, equation (11), $PM_{2,5}$ is the most influential variable with highest value of standardized coefficient, where higher levels of this pollutant are associated with increased CO_2 emissions. Equation (12) in the case of SVN, shows that $PT$is the most important variable and that higher PT values are associated with lower CO_2 emissions. Additionally, $GDPpc$ is positively associated with CO_2emissions. It is interesting to notice that for SVN with the increase of population CO_2emissions decrease.

Table 1 shows that the models have high performances, withhighR^2 values and low MSE. Among all statistically significant variables across the models, the VIF is low, indicating that there is no multicollinearity among the input variables.

Table 1.Summary of linear regression models for BIH, CRO and SVN

Variable	Country	St. Coef	$t-value$	$p-value$	VIF	R^2	MSE
REC	BIH	-0,317	-2,85	0,011	2,43	91,34%	0,319
UP		1,514	10,20	0,000	4,32		
FA		-0,551	-3,06	0,007	6,37		
$PM_{2,5}$	CRO	0,813	7,34	0,000	1	77,91%	0,495
FDI		0,335	3,03	0,007	1		
$GDPpc$	SVN	0,573	3,25	0,004	2,46	77,20%	0,503
PT		-1,240	-7,03	0,000	2,46		

3.2.Hypothesis tests for the differences between two means between each country

Hypothesis tests for the differences between the means of all variables between Bosnia and Herzegovina, Croatia, and Slovenia were conducted. Given that each variable consisted of 20 samples, $F-test$ was first performed to examine whether variances were different. Depending on the outcome of the $F-test$, either $t-test$ assuming equal variances or $t-test$ assuming unequal variances were applied to assess the differences in means. The proposed hypotheses for $F-test$ are shown in equations (13 – 15).

$$H_0: s_{BIH}^2 = s_{CRO}^2$$
$$1.2 \quad H_1: s_{BIH}^2 \neq s_{CRO}^2 \tag{13}$$

$$H_0: s_{BIH}^2 = s_{SVN}^2$$
$$H_1: s_{BIH}^2 \neq s_{SVN}^2 \tag{14}$$

$$H_0: s_{CRO}^2 = s_{SVN}^2$$
$$H_1: s_{CRO}^2 \neq s_{SVN}^2 \tag{15}$$

Tables 2 – 4 show results of $F-test$ for differences of variances between all countries, BIH – CRO, BIH – SVN, and CRO – SVN respectively, where Table 2, Table 3, and Table 4 show decisions for proposed hypotheses in equation (13), equation (14), and equation (15) respectively.

Table 2.F-test results – BIH vs. CRO

Variable	s_{BIH}^2	s_{CRO}^2	F_{stat}	$p-value$	F_{crit}	Decision
REC	47,81	11,51	4,155	0,001	2,124	Reject H_0
$PM_{2,5}$	14,84	7,02	2,114	0,051	2,124	Do not reject H_0
$GDPpc$	2,48·10^6	1,16·10^7	4,676	0,001	2,124	Reject H_0
FDE	1,45·10^{17}	1,71·10^{18}	11,779	0,000	2,124	Reject H_0
UP	4,12	1,5	2,755	0,014	2,124	Reject H_0
PT	9,89·10^{10}	2,05·10^{10}	4,825	0,000	2,124	Reject H_0
FA	114117,23	25102,21	4,546	0,001	2,124	Reject H_0
CO_2	13,37	4,94	2,709	0,015	2,124	Reject H_0

Table 3.F-test results – BIH vs. SVN

Variable	s_{BIH}^2	s_{SVN}^2	F_{stat}	$p-value$	F_{crit}	Decision
REC	47,81	3,68	12,980	0.000	2,124	Reject H_0
$PM_{2,5}$	14,84	5.92	2,507	0.023	2,124	Reject H_0
$GDPpc$	2,48·10^6	2,77·10^7	11,175	0.000	2,124	Reject H_0
FDE	1,45·10^{17}	4,67·10^{17}	3,215	0.006	2,124	Reject H_0
UP	4,12	1,95	2,115	0.051	2,124	Do not reject H_0
PT	9,89·10^{10}	1,24·10^9	79.563	0.000	2,124	Reject H_0
FA	114117,23	2284,15	49.961	0.000	2,124	Reject H_0
CO_2	13,37	2,02	6.620	0.000	2,124	Reject H_0

Table 4.F-test results – CRO vs. SVN

Variable	s^2_{CRO}	s^2_{SVN}	F_{stat}	$p-value$	F_{crit}	Decision
REC	11,51	3,68	3,124	0.007	2,124	Reject H_0
$PM_{2,5}$	7,02	5.92	1,186	0.353	2,124	Do not reject H_0
$GDPpc$	1,16·10^7	2,77·10^7	2,390	0.029	2,124	Reject H_0
FDE	1,71·10^{18}	4,67·10^{17}	3,664	0.003	2,124	Reject H_0
UP	1,5	1,95	1,302	0.280	2,124	Do not reject H_0
PT	2,05·10^{10}	1,24·10^9	16.490	0.000	2,124	Reject H_0
FA	25102,21	2284,15	10.990	0.000	2,124	Reject H_0
CO_2	4,94	2,02	2,444	0.026	2,124	Reject H_0

In Tables 2 – 4 where the decision was to 'Do not reject H_0', $t-test$ assuming equal variances was done, otherwise $t-test$ assuming non-equal variances was performed. From Tables 5 – 7 the results from performed $t-test$ between all countries can be seen.

From Table 5 it can be seen that there is enough statistical evidence to conclude that CRO has higher REC, $GDPpc$, FDI, UP, and PT, while BIH has higher $PM_{2,5}$ and FA. Also, there is not enough statistical evidence to conclude that BIH has higher CO_2emissions than CRO. Although the null hypothesis for CO_2 emissions was not rejected, as it was seen in equations (9) and (10) different independent variables are statistically significant in the final linear regression forCO_2 emissions predictions showing the complexity of this problem for each country individually.

Table 6 shows results of $t-test$ of differences between means of all variables for BIH and SVN. There is enough statistical evidence to conclude that BIH has higher $PM_{2,5}$, PT, FA and CO_2 emissions than SVN, while SVN has higher $GDPpc$, FDI, and UP than BIH. Also, there is not enough statistical evidence to conclude that BIH has higher REC than SVN.

Table 5.t-test results – BIH vs. CRO

Variable	μ_{BIH}	μ_{CRO}	Hypothesis	$t-stat$	$p-value$	$t-crit$	Decision
REC	21,63	28,59	$H_0: \mu_{CRO} - \mu_{BIH} \leq 0$ $H_1: \mu_{CRO} - \mu_{BIH} > 0$	4,139	0,000	1,699	Reject H_0
$PM_{2,5}$	34,97	21,45	$H_0: \mu_{BIH} - \mu_{CRO} \leq 0$ $H_1: \mu_{BIH} - \mu_{CRO} > 0$	13,256	0,000	1,684	RejectH_0
$GDPpc$	4080,87	12172,27	$H_0: \mu_{CRO} - \mu_{BIH} \leq 0$ $H_1: \mu_{CRO} - \mu_{BIH} > 0$	9,881	0,000	1,701	Reject H_0
FDI	$5,32 \cdot 10^8$	$2,08 \cdot 10^9$	$H_0: \mu_{CRO} - \mu_{BIH} \leq 0$ $H_1: \mu_{CRO} - \mu_{BIH} > 0$	5,211	0,000	1,714	Reject H_0
UP	45,60	55,28	$H_0: \mu_{CRO} - \mu_{BIH} \leq 0$ $H_1: \mu_{CRO} - \mu_{BIH} > 0$	18,720	0,000	1,692	Reject H_0
PT	$3,80 \cdot 10^6$	$4,22 \cdot 10^6$	$H_0: \mu_{CRO} - \mu_{BIH} \leq 0$ $H_1: \mu_{CRO} - \mu_{BIH} > 0$	5,654	0,000	1,701	Reject H_0
FA	21330,32	19142,41	$H_0: \mu_{BIH} - \mu_{CRO} \leq 0$ $H_1: \mu_{BIH} - \mu_{CRO} > 0$	26,871	0,000	1,701	Reject H_0
CO_2	20,31	20,04	$H_0: \mu_{BIH} - \mu_{CRO} \leq 0$ $H_1: \mu_{BIH} - \mu_{CRO} > 0$	0,295	0,385	1,692	Do not reject H_0

Table 6. t-test results – BIH vs. SVN

Variable	μ_{BIH}	μ_{SVN}	Hypothesis	$t-stat$	$p-value$	$t-crit$	Decision
REC	21,63	19,543	$H_0: \mu_{BIH} - \mu_{SVN} \leq 0$ $H_1: \mu_{BIH} - \mu_{SVN} > 0$	1,332	0,098	1,714	Do not reject H_0
$PM_{2,5}$	34,97	19,130	$H_0: \mu_{BIH} - \mu_{SVN} \leq 0$ $H_1: \mu_{BIH} - \mu_{SVN} > 0$	15,933	0,000	1,691	RejectH_0
$GDPpc$	4080,87	20842,51	$H_0: \mu_{SVN} - \mu_{BIH} \leq 0$ $H_1: \mu_{SVN} - \mu_{BIH} > 0$	13,976	0,000	1,711	Reject H_0
FDI	$5,32 \cdot 10^8$	$9,04 \cdot 10^8$	$H_0: \mu_{SVN} - \mu_{BIH} \leq 0$ $H_1: \mu_{SVN} - \mu_{BIH} > 0$	2,182	0,018	1,696	Reject H_0
UP	45,60	52,70	$H_0: \mu_{SVN} - \mu_{BIH} \leq 0$ $H_1: \mu_{SVN} - \mu_{BIH} > 0$	13,221	0,000	1,684	Reject H_0
PT	$3,80 \cdot 10^6$	$2,04 \cdot 10^6$	$H_0: \mu_{BIH} - \mu_{SVN} \leq 0$ $H_1: \mu_{BIH} - \mu_{SVN} > 0$	25,453	0,000	1,721	Reject H_0
FA	21330,32	12422,61	$H_0: \mu_{BIH} - \mu_{SVN} \leq 0$ $H_1: \mu_{BIH} - \mu_{SVN} > 0$	119,646	0,000	1,721	Reject H_0
CO_2	20,31	15,91	$H_0: \mu_{BIH} - \mu_{SVN} \leq 0$ $H_1: \mu_{BIH} - \mu_{SVN} > 0$	5,133	0,000	1,706	Reject H_0

Table 7. t-test results – CRO vs. SVN

Variable	μ_{CRO}	μ_{SVN}	Hypothesis	$t-stat$	$p-value$	$t-crit$	Decision
REC	28.59	19,543	$H_0: \mu_{CRO} - \mu_{SVN} \leq 0$ $H_1: \mu_{CRO} - \mu_{SVN} > 0$	10,633	0,000	1,694	Reject H_0
$PM_{2,5}$	21,45	19,130	$H_0: \mu_{CRO} - \mu_{SVN} \leq 0$ $H_1: \mu_{CRO} - \mu_{SVN} > 0$	2,952	0,003	1,684	Reject H_0
$GDPpc$	12172,27	20842,51	$H_0: \mu_{SVN} - \mu_{CRO} \leq 0$ $H_1: \mu_{SVN} - \mu_{CRO} > 0$	6,336	0,000	1,691	Reject H_0
FDE	$2{,}08 \cdot 10^9$	$9{,}04 \cdot 10^8$	$H_0: \mu_{CRO} - \mu_{SVN} \leq 0$ $H_1: \mu_{CRO} - \mu_{SVN} > 0$	3,654	0,000	1,697	Reject H_0
UP	55,28	52,70	$H_0: \mu_{CRO} - \mu_{SVN} \leq 0$ $H_1: \mu_{CRO} - \mu_{SVN} > 0$	6,356	0,000	1,684	Reject H_0
PT	$4{,}22 \cdot 10^6$	$2{,}04 \cdot 10^6$	$H_0: \mu_{CRO} - \mu_{SVN} \leq 0$ $H_1: \mu_{CRO} - \mu_{SVN} > 0$	67,876	0,000	1,717	Reject H_0
FA	19142,41	12422,61	$H_0: \mu_{CRO} - \mu_{SVN} \leq 0$ $H_1: \mu_{CRO} - \mu_{SVN} > 0$	186,079	0,000	1,711	Reject H_0
CO_2	20,04	15,91	$H_0: \mu_{CRO} - \mu_{SVN} \leq 0$ $H_1: \mu_{CRO} - \mu_{SVN} > 0$	7,158	0,000	1,691	Reject H_0

Results of $t - test$ for differences between means for all variables for CRO and SVN are shown in Table 7, where there is enough statistical evidence to conclude that CRO has higher values of REC, $PM_{2,5}$, FDI, UP, PT, FA and CO_2 than SVN, while SVN has higher value of $GDPpc$ than CRO.
From these results it can be seen that SVN has significantly lower values of CO_2 emissions than both BIH and CRO showing that they are much further in the decarbonization process than BIH and CRO, while other variables vary differently between countries.

4. Conclusion

This research shows that CO_2emissions and other variables analyzed in this research vary significantly between Bosnia and Herzegovina, Croatia and Slovenia, which indicates the need for an individualized approach in creating decarbonization strategies for each country. Using machine learning method multiple linear regression with backward elimination, models were developed that allow the identification of variables with a statistically significant impact on CO_2 emissions.
For Bosnia and Herzegovina, developed linear regression model showed that increase of urban population increasesCO_2 emissions, while the increase of values of renewable energy consumption and forest area reducesCO_2 emissions. This clearly indicates the areas where Bosnia and Heregovina needs to focus its attention in order to achieve its decarbonization goals. The developed model for Croatia shows that $PM_{2,5}$air pollution was the most influential variable, where the increase of $PM_{2,5}$ increases values of CO_2 emissions, suggesting a need to focus on measures to improve air quality. In the case of Slovenia, the developed model shows that with the increase of total population the value of CO_2 emissions decreases, which may indicate a higher level of environmental awareness, while the increase of GDP per capitaincreases CO_2 emissions.
In addition to the models, conducted hypothesis tests show that there are significant differences in the means of the most variables between the countries, which further confirms the complexity of this problem and the necessity of locally adapted measures. However, it is important to emphasize results show that there is enough statistical evidence to conclude that Slovenia has significantly lower mean values of CO_2 emissions compared to Bosnia and Herzegovina and Croatia, while there is not enough statistical evidence to conclude that CO_2 emissionsare higher in BiH than in Croatia. These results show that Slovenia was more successful in reducing CO_2 for the analyzed period. This research shows the potential of using machine learning in data analysis and decision-making support in the field of sustainable development and CO_2 emission. Future research recommends using other machine learning

algorithms and to include other variables that may influence CO_2 emissions in order to support the transition to sustainable industrial practices which are highly important in transitioning to Industry 5.0.

5. References

[1] UN Environment Programme, "How do greenhouse gases actually warm the planet?," UN Environment Programme Web site, 5 January 2022, [Online]. Available: https://www.unep.org/news-and-stories/story/how-do-greenhouse-gases-actually-warm-planet. [Accessed 05 April 2025].

[2] Council of the European Union, "Preparations for the 29th Conference of the Parties (COP29) of the United Nations Framework Convention on Climate Change (UNFCCC): Council conclusions," Council of the European Union, Brussels, 2024.

[3] UNFCCC, "Key aspects of the Paris Agreement," UNFCCC, [Online]. Available: https://unfccc.int/most-requested/key-aspects-of-the-paris-agreement. [Accessed 05 February 2025].

[4] United Nations Climate Action, "For a livable climate: Net-zero commitments must be backed by credible action," United Nations Climate Action Web site, [Online]. Available:https://www.un.org/en/climatechange/net-zero-coalition. [Accessed 05 April 2025].

[5] European Commission, "Delivering the European Green Deal," European Commission Web site, [Online]. Available:https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-greendeal/delivering-european-green-deal_en#boosting-global-climate-action. [Accessed 05 April 2025].

[6] European Commission, "About the EU ETS," European Commission Web site, [Online]. Available: https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/about-eu-ets_en#what-is-the-eu-ets. [Accessed 05 April 2025].

[7] European Commission, "Carbon Border Adjustment Mechanism," 25 March 2025. [Online]. Available: https://taxation-customs.ec.europa.eu/carbon-border-adjustment-mechanism_en#cbam-transitional-phase-2023--2025. [Accessed 05 April 2025].

[8] H. Ritchie, P. Rosado and M. Roser, "Greenhouse gas emissions," Our World in Data, June 2020. [Online]. Available: https://ourworldindata.org/greenhouse-gas-emissions. [Accessed 05 April 2025].

[9] H. Ritchie and M. Roser, "CO_2 emissions," Our World in Data, January 2024, [Online]. Available: https://ourworldindata.org/co2-emissions.

[Accessed 05 April 2025].

[10] IEA, “The relationship between growth in GDP and CO2 has loosened; it needs to be cut completely,” IEA, Paris, 2024, [Online]. Available: https://www.iea.org/commentaries/the-relationship-between-growth-in-gdp-and-co2-has-loosened-it-needs-to-be-cut-completely. [Accessed 05 April 2025].

[11] I. Jianu, S. Jeloaica and M. Tudorache, "Greenhouse Gas Emissions and its Main Drivers: a Panel Assessment for EU-27 Member States," American International Journal of Business Management, vol. 5, no. 5, pp. 138-146, 2022. arXiv:2205.00295

[12] A. A. A. Chowdhury, A. H. Rafi, A. Sultana and A. All Noman, "Enhancing Green Economy with Artificial Intelligence: Role of Energy Use and FDI in the United States," Journal of Environmental and Energy Economics, pp. 55–76, 2024. arXiv:2501.14747.

[13] L. P. Fávero, R. D. F. Souza, P. Belfiore, M. R. Luppe and M. Severo, "Global relationship between economic growth and CO2 emissions across time: a multilevel approach," International Journal of Global Warming, vol. 26, no. 1, pp. 38-63, 2021. doi:10.1504/IJGW.2022.120067.

[14] B. P. Chukwunonso *et al.*, “Predicting carbon dioxide emissions in the United States of America using machine learning algorithms,” Environ Sci Pollut Res, vol. 31, no. 23, pp. 33685–33707, 2024, doi: 10.1007/s11356-024-33460-1.

[15] L. Farahzadi and M. Kioumarsi, “Application of machine learning initiatives and intelligent perspectives for CO2 emissions reduction in construction,” Journal of Cleaner Production, vol. 384, p. 135504, 2023, doi: 10.1016/j.jclepro.2022.135504.

[16] Y. Li and Y. Sun, “Modeling and predicting city-level CO2 emissions using open access data and machine learning,” Environ Sci Pollut Res, vol. 28, no. 15, pp. 19260–19271, 2021, doi: 10.1007/s11356-020-12294-7.

[17] Y. Meng and H. Noman, “Predicting CO2 Emission Footprint Using AI through Machine Learning,” Atmosphere, vol. 13, no. 11, p. 1871, 2022, doi: 10.3390/atmos13111871.

[18] S. Li, Y. W. Siu and G. Zhao, "Driving Factors of CO2 Emissions: Further Study Based on Machine Learning," Frontiers in Environmental Science, vol. 9, 2021, doi: 10.3389/fenvs.2021.721517.

[19] C. Magazzino, M. Mele, and N. Schneider, “A machine learning approach on the relationship among solar and wind energy production, coal consumption, GDP, and CO2 emissions,” Renewable Energy, vol. 167, pp. 99–115, 2021, doi: 10.1016/j.renene.2020.11.050.

[20] H. Jin et al., “Computational simulation using machine learning models in prediction of CO2 absorption in environmental applications,” Journal of Molecular Liquids, vol. 358, p. 119159, 2022, doi: 10.1016/j.molliq.2022.119159.
[21] S. Asumadu-Sarkodie and P. A. Owusu, "Recent evidence of the relationship between carbon dioxide emissions, energy use, GDP, and population in Ghana: A linear regression approach," Energy Sources, Part B: Economics, Planning, and Policy, vol. 12, no. 6, pp. 495-503, 2017. doi: 10.1080/15567249.2016.1208304.
[22] C. S. Choi and L. Abdullah, "Prediction of Carbon Dioxide Emissions Using Two Linear Regression-based Models: A Comparative Analysis," Journal of Applied Engineering (JOAE), vol. 4, no. 4, 2016.
[23] H. Horner, "The Shift from Industry 4,0 to 5.0: A Global Perspective on the Future of Manufacturing," Engineerging Institute of Technology (EIT), 11 April 2025. [Online]. Available: https://www.eit.edu.au/industry-4-0-to-5-0-perspective-on-manufacturing/. [Accessed 12 April 2025].
[24] M. Ghobakhloo, M. Iranmanesh, M. F. Mubarak, M. Mubarik, A. Rejeb, and M. Nilashi, “Identifying industry 5.0 contributions to sustainable development: A strategy roadmap for delivering sustainability values,” Sustainable Production and Consumption, vol. 33, pp. 716–737, 2022, doi: 10.1016/j.spc.2022,08.003.
[25] C. K. K. Reddy, P. R. Anisha, S. Khan, M. M. Hanafiah, L. Pamulaparty, and R. M. Mohana, Eds., Sustainability in industry 5.0: theory and applications, First edition. in Computational methods for industrial applications. Boca Raton London New York: CRC Press, 2024.
[26] S. Aheleroff, H. Huang, X. Xu, and R. Y. Zhong, “Toward sustainability and resilience with Industry 4,0 and Industry 5.0,” Front. Manuf. Technol., vol. 2, pp. 951643, 2022, doi: 10.3389/fmtec.2022,951643,
[27] S. Bamdad, “Leveraging machine learning and decision analytics for sustainable and resilient environmental monitoring in metal processing industries: a step towards Industry 5.0,” International Journal of Production Research, pp. 1–27, 2025, doi: 10.1080/00207543.2025.2487567.
[28] B. Martini, D. Bellisario, and P. Coletti, “Human-Centered and Sustainable Artificial Intelligence in Industry 5.0: Challenges and Perspectives,” Sustainability, vol. 16, no. 13, p. 5448, 2024, doi: 10.3390/su16135448.
[29] "World Bank Open Data," [Online]. Available: https://data.worldbank.org. [Accessed 05 April 2025].
[30] I. Tom, Statistics Plain and Simple, Second Edition. Kendall/Hunt Publishing Co,U.S., 2013.