Proceedings of the 3rd International Conference on Computing Innovation and Applied Physics
DOI: 10.54254/2753-8818/25/20240898

Analysis of influencing factors of carbon emissions in China based on the STIRPAT Model

Wenqing Mao

Department of Mathematics, King's College London, London, United Kingdom.

K22019633@kcl.ac.uk

Abstract. China, as a major economic power, has been increasing its carbon emissions year after year. Effectively controlling carbon emissions and finding suitable and effective methods to reduce emissions have become the main research themes of current research. The Stochastic Impacts by Regression on Population, Affluence, and Technology (STIRPAT) model is used in this work to analyze the impact of GDP, population, urbanization, and energy intensity on China's carbon emissions from 2003 to 2020. From the output by the SPSS software, it can be illustrated that GDP and energy intensity have more obvious contribution on carbon emission, while urbanization level and population don't. Additionally, as the GDP index increases by a value of one, a 1.220 change will be seen by the carbon emission. Similarly, every one unit change for energy intensity is associated with 0.897 change in carbon emission. Therefore, this paper can consider effective ways to conserve energy and mitigate greenhouse gas emissions from these two aspects, and in this way attain the objective of carbon peaking and carbon neutrality.

Keywords: Carbon Emission, STIRPAT Model, Regression, China.

1. Introduction

Carbon emissions represent the release of carbon dioxide gas into the atmosphere because of a variety of anthropogenic activities. Due to the fact that carbon emissions make up a sizable portion of greenhouse gas emissions, which cause climate change and global warming, and are closely linked to sustainable development, it is crucial to pay attention to it.

First, one of the primary causes of climate change is carbon emissions. The greenhouse effect is intensified in the Earth's atmosphere as a result of greenhouse gas emissions, such as carbon dioxide, which in turn triggers repercussions of climate change, including severe weather, sea level rise, floods, and droughts [1]. This poses a threat to the global society and economy and calls for action to slow down the rate of climate change. Secondly, carbon emissions also have negative impacts on ecosystems and biodiversity. Rising temperatures, acid rain, and air pollution directly affect plants, animals, and habitats, accelerating the loss of biodiversity and disrupting the ecological balance. In addition, carbon emissions are related to sustainable development [2]. Excessive carbon emissions led to energy wastage and resource depletion, which may trigger energy crises and social instability. Therefore, focusing on carbon emissions is not only key to combating climate change but also involves many aspects of ecological protection, sustainable development, and international cooperation, which are essential to safeguarding the future of the planet and human society.

© 2024 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/25/20240898

It is particularly important to study which development model can be adopted to mitigate carbon emissions. It is a widely recognized fact that energy consumption is a major factor in the growth of carbon emissions [3]. But what is the relationship between carbon dioxide emissions and factors such as a country's population size, affluence, technology, and level of urbanization? To solve this problem, the exports established a model called the Stochastic Impacts by Regression on Population, Affluence, and Technology (STIRPAT) model [4]. It is a statistical model used to examine and forecast the effects of human activity on the environment in environmental and ecological studies. The model was developed as a tool to understand the driving forces behind environmental changes, such as deforestation, carbon emissions, or biodiversity loss [4].

In this essay, the STIRPAT model is used to quantitatively analyze the relation between carbon emission and population, GDP, energy intensity, and urbanization level.

2. Methodology

2.1. Introduction to the model

The STIRPAT model is developed from the IPAT model, which is a widely acknowledged formula for evaluating the impact of human activities on the environment. I = (PAT) is a mathematical symbol used to describe a formula for the impact of human activity on the environment. Three factors—population (P), wealth (A), and technology (T)—determine how much of an impact humans have on the environment. It resembles the Kaya identity, which relates particularly to carbon dioxide emissions, in terms of form [5].

The STIRPAT model is a more advanced and quantitative tool that allows for empirical analysis and precise estimation of these relationships between population, affluence, and technology as drivers of environmental impact [5]. It is commonly used for research that requires data-driven assessments of environmental impacts. It builds upon the IPAT concept but incorporates statistical techniques to analyze data and estimate coefficients [5].

The STIRPAT model is rooted in the idea that environmental impact (Y) can be expressed as a function of multiple factors:

$$Y = f(P, A, T, X_1, X_2 \ldots, X_n) \tag{1}$$

Where:

- **Y** represents the environmental impact or outcome of interest.
- **P** stands for population, indicating the size and demographic characteristics of the human population.
- **A** represents affluence, which measures economic prosperity or consumption patterns.
- **T** denotes technology and accounts for technological factors that affect the environment.
- $\mathbf{X_1, X_2, \ldots X_n}$ represent additional independent variables that can influence the environmental impact. These variables can include social, economic, political, and cultural factors, depending on the research focus [4, 5].

The STIRPAT model posits that these factors interact in a multiplicative way, and the relationships are estimated through statistical regression analysis. Researchers use historical data and statistical techniques to estimate the model's coefficients, which shed light on the potency and direction of the relationships between the variables. The model can be used to make anticipations, test hypotheses, and assess the relative importance of different drivers of environmental impact.

2.2. Variable settings

The general formula for this model is shown in Formula 2:

$$I = aP^bA^cT^d \tag{2}$$

For which a is a constant and b, c, and d are parameters of each influencing factor. In this model, the dependent variable represents the environmental outcome, which is the carbon emissions. The independent variables are P, A and T [6, 7]. P represents population, which is the size of the population in the area or region of interest. It represents the number of people and their demographic characteristics.

A represents affluence, a measure of economic prosperity or consumption patterns. It is often represented by variables such as Gross Domestic Product, income levels, or other indicators of wealth. T represents technology [7]. This variable accounts for technological factors that influence environmental impact. It can include measures of technological advancement, energy efficiency, or the adoption of cleaner technologies.

Based on this model and combined with the regional characters of China, I choose population, GDP, energy intensity, and urbanization level as the independent variables and carbon emission as the dependent variable [8].

2.3. Data source and processing

The data used in this essay are from the Climate Watch website and national data from the National Bureau of Statistics [9, 10]. The graph below shows the changes in carbon emissions for the last thirty years [9]. As shown in Figure 1below, the carbon emission of China increased significantly for twenty years from 1990 and slightly raised for the last ten years. This reflects that controlling carbon emissions is a vital problem.

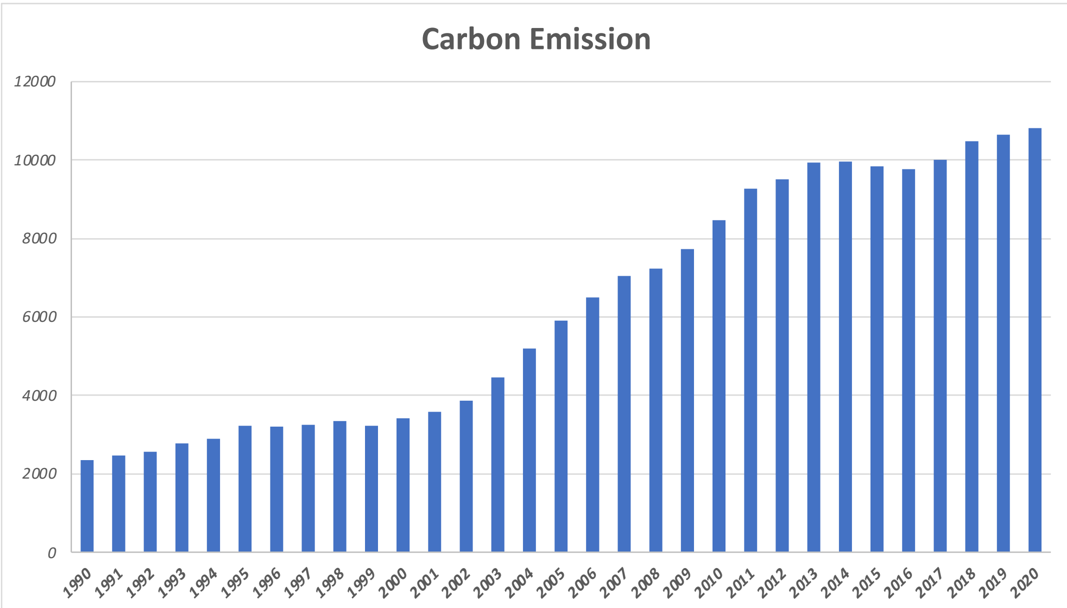

Figure 1. Carbon Emission of China from 1990 to 2020 (Photo/Picture credit: Original)

Data of the independent variables are from the national website of China. It is shown in Table 1 below, about the population, people in urban areas, GDP, and carbon consumption [10].

Table 1. Population, People in Urban Areas, GDP, and Carbon Consumption of China for Last Twenty Years

	Population	People in urban area	GDP	Carbon Consumption
2003	129227	52376	137422	197083
2004	129988	54283	161840.2	230281
2005	130756	56212	187318.9	261369
2006	131448	58288	219438.5	286467
2007	132129	60633	270092.3	311442
2008	132802	62403	319244.6	320611
2009	133450	64512	348517.7	336126

DOI: 10.54254/2753-8818/25/20240898

Table 2. (continued)

2010	134091	66978	412119.3	360648
2011	134916	69927	487940.2	387043
2012	135922	72175	538580	402138
2013	136726	74502	592963.2	416913
2014	137646	76738	643563.1	428334
2015	138326	79302	688858.2	434113
2016	139232	81924	746395.1	441492
2017	140011	84343	832035.9	455827
2018	140541	86433	919281.1	471925
2019	141008	88426	986515.2	487488
2020	141212	90220	1013567	498314
2021	141260	91425	1149237	525896

To find out the urbanization level and energy intensity, data processing before establishing the model is needed. The percentage of the people living in urban areas in relation to the total population is a common way to gauge the degree of urbanisation in a region or nation [11]. The equation is Formula 3,

$$U = \frac{people\ in\ urban\ areas}{population} \tag{3}$$

The calculated urbanization level is shown in Table 2 below in the graph.

Table 3. Calculated Urbanization Level

Year	U	Year	U
2003	0.4053	2013	0.5449
2004	0.4176	2014	0.5575
2005	0.4299	2015	0.5733
2006	0.4434	2016	0.5884
2007	0.4589	2017	0.6024
2008	0.4699	2018	0.615
2009	0.4834	2019	0.6271
2010	0.4995	2020	0.6389
2011	0.5183	2021	0.6472
2012	0.531	2022	0.6522

In addition, energy intensity is a metric for measuring the effectiveness of energy use and is typically quantified in terms of the amount of energy expended for each unit of economic output or other relevant metrics [11]. Its equation is defined by Formula 4,

$$I = \frac{Total\ Energy\ Consumption}{Gross\ Domestic\ Product} \tag{4}$$

The calculated energy intensity from 2003 to 2021 is indicated in the Table 3.

DOI: 10.54254/2753-8818/25/20240898

Table 4. Calculated Energy Intensity

Year	I	Year	I
2003	1.4341	2013	0.7031
2004	1.4229	2014	0.6656
2005	1.3953	2015	0.6302
2006	1.3055	2016	0.5915
2007	1.1531	2017	0.5478
2008	1.0043	2018	0.5134
2009	0.9644	2019	0.4942
2010	0.8751	2020	0.4916
2011	0.7932	2021	0.4576
2012	0.7467		

From the above analysis, it is known that the relationship between environmental effects and their drivers is $\mathrm{I} = \mathrm{aP^bA^cT^de}$. By taking logarithm to both sides, it can get formula 5,

$$\ln I = \ln a + b \ln P + c \ln A + d \ln T + \ln e \tag{5}$$

By taking logarithms, these order-of-magnitude differences can be reduced, making the relative impacts between variables easier to compare and understand [11]. At the same time, taking logarithms transforms the multiplicative relationship into an additive one, making the model easier to analyze with linear regression. The processed data is illustrated below in Table 4.

Table 5. Data of Population, GDP, Urbanization, Energy Intensity and Carbon Consumption after Data Processing for the Last Twenty Years

	Population	GDP	Urbanization	Energy Intensity	Carbon Consumption
2003	11.7693258	11.8308118	-0.9031221	0.36056848	8.40351821
2004	11.7751974	11.9943647	-0.873231	0.35269087	8.55382787
2005	11.7810883	12.1405678	-0.8442027	0.33312069	8.6828483
2006	11.7863666	12.2988273	-0.8132151	0.26655134	8.77835211
2007	11.791534	12.506519	-0.7789394	0.14244937	8.85899975
2008	11.7966146	12.6737129	-0.7552459	0.00427097	8.88758779
2009	11.8014822	12.7614443	-0.7268756	-0.0362029	8.95292753
2010	11.806274	12.9290681	-0.6941545	-0.1334105	9.04522513
2011	11.8124076	13.0979481	-0.6572005	-0.2316571	9.1341157
2012	11.8198365	13.1966913	-0.6329875	-0.2921407	9.16078872
2013	11.8257342	13.2928876	-0.607153	-0.3522548	9.20454561
2014	11.8324405	13.3747754	-0.5842881	-0.4071168	9.206497
2015	11.8373685	13.4427907	-0.5563499	-0.4617306	9.19488657
2016	11.8438969	13.5230104	-0.5303496	-0.5250952	9.18808454
2017	11.8494763	13.6316309	-0.5068292	-0.6017622	9.21077028
2018	11.8532545	13.7313472	-0.4861297	-0.6667719	9.25737089
2019	11.8565719	13.801934	-0.4666506	-0.7049131	9.27397129
2020	11.8580176	13.8289864	-0.4480112	-0.7100007	9.28903232

DOI: 10.54254/2753-8818/25/20240898

2.4. Result analysis

Multicollinearity is a correlation between multiple independent variables that can render the coefficients in the model impractical when performing regression calculations. By using SPSS (Statistical Package for the Social Sciences) data analysis software to test the covariance of the relevant independent variables, it is shown in Table 5 that the four factors, population, GDP, urbanization, and energy intensity do not have collinear relationships.

Table 6. Collinearity Diagnotics of the Listed Influencing Factors

Dimension	Eigenvalue	Condition Index	(Constant)	Populat ion	GDP	Urbaniz ation	Energy Intensity
1	4.207	1.000	0.00	0.00	0.00	0.00	0.00
2	0.792	2.304	0.00	0.00	0.00	0.00	0.01
3	0.000	122.893	0.00	0.00	0.01	0.09	0.95
4	1.336E-5	560.784	0.00	0.00	0.45	0.04	0.04
5	5.534E-9	27572.233	1.00	1.00	0.55	0.87	0.01

R-squared (R^2), otherwise called the coefficient of determination, is employed to gauge the degree to which the dependent variables' variance is explained by the independent variables in a regression model [12]. It quantifies the goodness of fit of the model to the observed data. As can be seen in Table 6 below, it is the result of the R square generated by SPSS software. It can be deduced that the R square is 1.000, which means that the linear regression model is fit for it in this situation. It shows that there is a linear model between the independent and dependent variables with a good fit.

Table 7. R Square of the Model Output

R	R Square	Adjusted R Square	Std. Error of the Estinmate	R Square Change
0.998[a]	0.997	0.995	0.018	0.997

[a]Dependent Variable: Carbon Consumption

This table determines whether the model is critical for determining the consequences. It is depicted below in Table 7.

P-value/ Sig value: Generally, 95% confidence interval is selected for research, so that the p-value should be below 0.05. In Table 7, it is less than .001. Therefore, the result is significant [13].

F-ratio: It shows how well the model fits the data after accounting for the model's flaws, improving the variable's forecast. the model is valid if the F-value is greater than 1. In the above table, the value is 930.602, which means the model is sufficient [13].

Table 8. ANOVA Table of the Output

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1.204	4	0.301	930.602	<0.001
Residual	0.004	13	0.000	0.000	0.000
Total	1.208	17	0.301	930.602	<0.001

As a result, read from Table 8, the relationship equation between the correlation between conservation of energy and its influencing factors is

$$\ln Y = -0.524 \ln P + 2.944 \ln A - 0.183 \ln U + 1.279 \ln I \quad (6)$$

From Table 8 shown below, the relationship between carbon emission and each factor separately can be interpreted [13]. By reading the significant values associated with population and urbanization, both values are over 0.05, which means that they both don't have significant contribution to carbon emission. Instead, both significant value of GDP and energy intensity are less than 0.05, this points out that they both have significant contribution to carbon emission.

In addition, as the GDP index increases by a value of one, a 1.220 change will be seen by the carbon emission. Similarly, every one unit change for energy intensity is associated with 0.897 change in carbon emission. Therefore, this paper can consider effective ways to save energy and reduce emissions from these two aspects.

Table 9. Coefficient Output of the Model by Software

	(Constant)	Population	GDP	Urbanization	Energy Intensity
Unstandardized B	49.152	-4.743	1.220	-0.334	0.897
Coefficients Std.Error	41.083	3.340	0.102	1.044	0.135
Standardized Coefficients Beta	0.000	-0.524	2.944	-0.183	1.279
t	1.196	-1.420	11.936	-0.320	6.640
Sig.	0.253	0.179	<0.001	0.754	<0.001
Zero-order	0.000	0.931	0.973	0.948	0.943
Correlations partial	0.000	-0.366	0.957	-0.088	0.879
Part	0.000	-0.023	0.195	-0.005	0.109
Collinearity Tolerance	0.000	0.002	0.004	0.001	0.007
Statistics VIF	0.000	509.191	227.256	1224.858	138.648

3. Conclusion

This paper outlines the characteristics of energy consumption in China from 2003 to 2020, estimates CO_2 emissions and their intensity, and quantitatively studies how CO_2 emissions are affected by factors including population, GDP, urbanization, and energy use. based on the STIRPAT model combined with liner regression methods under the usage of SPSS software.

The results show that:

(1) from 2003 to 2020, China's overall energy use and CO_2 emissions have increased each year. Carbon dioxide emission intensity has generally maintained a decreasing trend, and China has gradually deepened its understanding of the carbon emission problem and taken a series of measures to reduce the intensity of carbon emission.

(2) Numerous variables have an impact on carbon dioxide emissions, including population, GDP, urbanization level, and energy intensity. The intensity of the effects of various variables on carbon emissions varies substantially. GDP is the most significant factor, while other factors have relatively small effects, with energy intensity, urbanization, and population in descending order of significance.

This shows out that low-carbon technologies have great potential for energy saving and emission reduction. Low-carbon technological innovation should be vigorously promoted and the proportion of low-carbon technological inputs in scientific and technological inputs should be enhanced; The government should reinforce its support and leadership for institutions and corporations to carry out

DOI: 10.54254/2753-8818/25/20240898

technological innovation in energy savings and emission reduction, and as a result, advance China's progress in reducing carbon emissions.

References

[1] Mohajan, H. 2013, Global greenhouse gas emissions and climate change, (LAP LAMBERT Academic Publishing).
[2] Ansuategi, A. Escapa, M. 2002, Economic growth and greenhouse gas emissions, (Ecological Economics, vol. 40), no. 1, pp: 23-37.
[3] Nejat, P. Jomehzadeh, F. Taheri, M. 2015, A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries), (Renewable and sustainable energy reviews, vol. 43), pp. 843-862.
[4] Fan, Y. Liu, L C. Wu, G. et al. 2006, Analyzing impact factors of CO2 emissions using the STIRPAT model, (Environmental Impact Assessment Review, vol. 26), no. 4, pp. 377-395.
[5] Bargaoui, S A. Liouane, N, Nouri, F Z. 2004, Environmental impact determinants: An empirical analysis based on the STIRPAT model, (Procedia-Social and Behavioral Sciences, vol. 109), pp. 449-458.
[6] Tang, D. Ma, T. Li, Z. et al. 2016, Trend prediction and decomposed driving factors of carbon emissions in Jiangsu Province during 2015–2020. (Sustainability, vol. 8), no. 10, pp. 1018.
[7] Niu, L. Zhagn, L. Xi, F. et al. 2023, Influencing factors and scenario forecasting of carbon emissions in Liaoning Province, China. (The Journal of Applied Ecology, vol. 34), no. 2, pp. 499-509.
[8] Yang, L. Xia, H. Zhang, X. et al. 2018, What matters for carbon emissions in regional sectors? A China study of extended STIRPAT model. (Journal of Cleaner Production, vol. 180), pp. 595-602.
[9] Historical GHG Emissions, September 21, 2023, https://www.climatewatchdata.org/ghg-emissions?breakBy=sector&chartType=line&end_year=2020&regions=TOP&sectors=energy,building,electricity-heat,fugitive-emissions,manufacturing-construction,other-fuel-combustion,transportation&start_year=1990.
[10] National Bureau of Statistics, September 21, 2023, https://data.stats.gov.cn/easyquery.htm?cn=C01.
[11] Yao, X. Zhou, H. Zhang, A. et al. 2015, Regional energy efficiency, carbon emission performance and technology gaps in China: A meta-frontier non-radial directional distance function analysis, (Energy Policy, vol. 84), pp. 142-154.
[12] Andrian, T. Sudibyo, Y A. 2019, Disclosure effect of carbon emission and corporate social responsibility to financial performance. (Journal of Economics and Sustainable Development, 12).
[13] Kafle, S C. 2019, Correlation and regression analysis using SPSS. (Management, Technology & Social Sciences, 126).