Analysing carbon dioxide emissions in the UK --by some explanatory models

Introduction

CO2 emissions

Carbon dioxide is one of the major contributors to Global Warming and air pollution. Global carbon dioxide emissions surged around 90% since 1970, while fossil fuel and industrial emissions increased from 1950 onwards (Ram C. Kafle et al., 2019). The UK is one of the largest sources of carbon emissions and may also suffer from serious pressure to reduce carbon dioxide emissions.

The Environmental Kuznets Curve (EKC) in Figure 1 was discovered in 1992. It reveals the connection between sulfur dioxide levels and the income per person (GDP) in 47 cities across 31 countries (Yandle et al., 2004). Initially, when the GDP increases, more greenhouse gases are released. However, as the economy shifts from industrial production to service-based industries, the environmental damage gradually decreases.

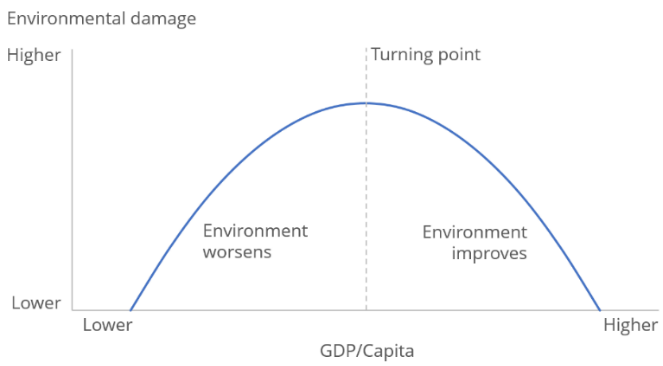

Figure 1. The Environmental Kuznets Curve (Source: https://www.ons.gov.uk)

Meanwhile, if the economy can grow while reducing carbon dioxide emissions, carbon dioxide emissions per unit of GDP (known as CO2 intensity) will need to fall (Hannesson R, 2020). Therefore, if carbon intensity is low, less CO2 is released to produce a certain amount of economic output, indicating that the economy is more carbon efficient.

Nowadays, there are many quantitative methods and factors to predict CO2 emissions, such as energy consumption, GDP, population and vehicular. Sun, D. et al (2023) have investigated the role of battery electric vehicles (BEVs) in reducing environmental pollution through an econometric approach. This study explores the correlation between gross domestic product (GDP), population, urbanization, BEVs, renewable energy consumption (REC), and carbon dioxide (CO2) in the United States, China, France, Germany, and Norway. What’s more, Ram C. Kafle expects to better understand the instantaneous rate of change of carbon emissions in the time domain by Differential equations. (Ram C. Kafle et al., 2019). They model the variation in CO2 emissions trend through data-driven differential equations.

Research Hypothesis

Therefore, although the above research has significant implications for policymakers, this study only includes the five countries without the UK.

Overall, my study aims to provide further insights and grounds for predicting and reducing CO2 emissions in the UK. That means I would follow the five independent variables (GDP, population, urbanization, BEVs and REC) that were mentioned above, searching out answers to the research question: "Is there a significant correlation between the five independent variables and CO2 emissions in the UK?". Then, I would choose London as an example which is the capital city in the UK, pointing out the higher carbon-efficient economy areas (showing low CO2 emissions and high economic contribution) and lower performance areas with high environmental costs.

Analysis

Variable and Data

The data for carbon emissions, gross domestic product, population, urbanization and renewable energy data between 1990 and 2020 in the UK were generated from WDI. Battery electric vehicles data spanning 2010 to 2020 was also gained from WDI. The periods were chosen according to data availability because the electric vehicle is still an emerging item. What's more, carbon emissions and gross domestic product data (2019) by borough in London were collected by the London datastore. Table 1 illustrates the measuring system, symbols and sources of data for the above data sets. Besides, those variables were displayed as line charts with years as x and the variable values as y in_ Figure 2.

Table 1. Variable and data source descriptions

Variable	Description	Source
Carbon dioxide	CO2 emissions from fossil fuel (kt) in the UK	WDI
Gross domestic product	the total value of all goods and services in the UK	WDI
Population	the total number of people living in the UK	WDI
Urbanization	Urban population in the UK (% of the entire population)	WDI
Electric vehicles (BEV)	The total of registered BEV in the UK	WDI
Renewable energy consumption	% of total renewable energy consumption in the UK	WDI
Carbon dioxide in London	CO2 emissions by boroughs in London	London datastore
Gross domestic product in London	the total value of all goods and services by boroughs in London	London datastore

Figure 2. Line charts of original data

Correlation matrix

Before modeling, we may want to know the correlation between variables. Therefore, we have obtained a correlation matrix for exploring the relation between Carbon dioxide emissions and our independent variables in Figure 3.

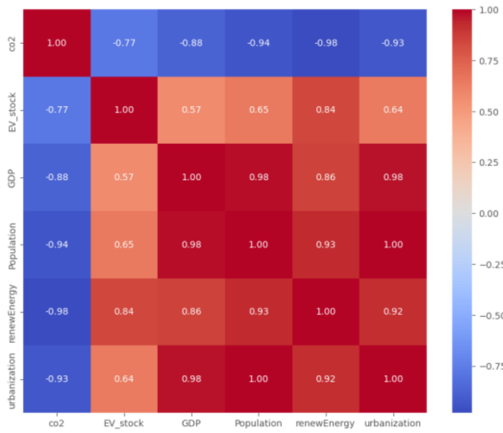

Figure 3. Correlation matrix

This matrix illustrates high relations between Carbon dioxide emissions and our independent variables. AS UK carbon dioxide (CO_2) emissions peaked in 1972(Dr Amina Syed, 2019), higher GDP means lower carbon emissions. The Pearson correlation coefficient shows a stronger negative value (-0.88) between CO_2 emissions and GDP. Meanwhile, independent variables may have strong multicollinearity with each other, especially population and relation with GDP and urbanization showing Pearson correlation coefficients of 0.98 and 1 respectively.

Simple linear regression

Before running a linear regression, I need to standardize our variables as it ensures that the coefficients of the regression results are meaningful.

The first explanatory variable examined is the number of gross domestic product (GDP). As we would like to know the relation between GDP as the independent variable and CO2 emissions as the dependent variable, we performed the simple linear regression using the OLS method in Figure 4.

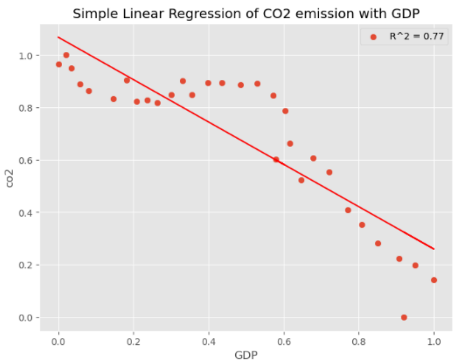

Figure 4. Simple Linear Regression of CO2 emissions with GDP

The result illustrates the negative trend with a coefficient of -0.81, while the R square is 0.77. Meanwhile, the P-value is almost equal to zero which means we can reject the zero hypothesis and believe that this model can predict changes in CO2 emissions. Then, we may think that the higher population and high-level urbanization may related to higher CO2 emissions respectively. To examine these relationships, we have plotted population and urbanization showing their relations with the CO2 emissions as well as building the simple linear regression models (Figure 5). The negative trends can be seen from the results with coefficients of -0.81 and -0.78 respectively, while the R squares climb to 0.88 and 0.87 respectively.

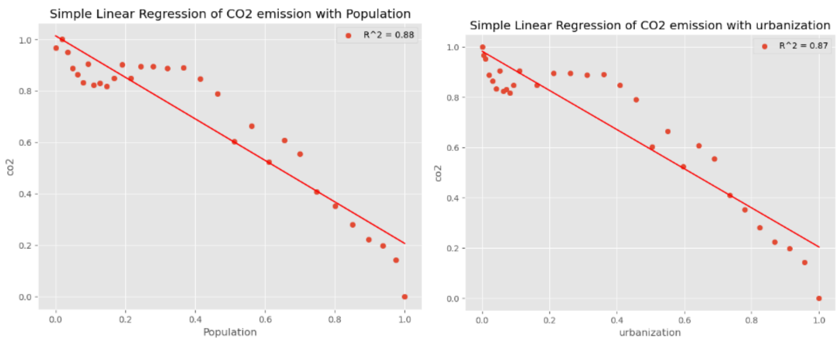

Figure 5. Simple Linear Regression of CO2 Emissions with Population and Urbanization

Finally, we consider that increasing the number of electric vehicles and applying renewable energy can mitigate carbon dioxide emissions because using electric cars powered by electricity would not produce air pollution. Some researchers have also proved that renewable energy utilization can reduce CO2 emissions in the atmosphere which has positive effects on the environment. (Zafar et al. 2020). Therefore, two single-linear regressions will be performed to check our above hypothesis. We then obtained the following results in Figure 6: The coefficients of renewable energy and EV cars are -0.98 and -0.9 respectively. Besides, the R-squares of the regressions are 0.96 and 0.72 respectively, while the P-values are lower than 0.001, showing significant results of the linear models. The results show that all independent variables are negatively correlated with CO2 emissions, while the strongest negative correlation is between renewable energy and CO2 emissions.

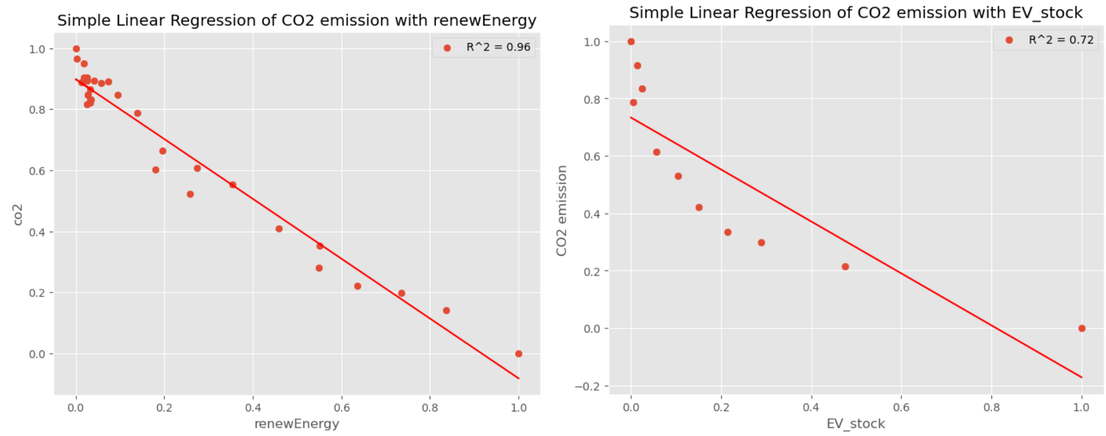

Figure 6. Simple Linear Regression of CO2 emissions with Renewable energy and Electricity cars

Although all of the independent variables may have strong relations with the CO2 emissions relied on the above regression results, differential equations can help us see how the response variable changes over time which otherwise can only be studied indirectly (Ram C. Kafle et al., 2019). It is more interesting to understand how fast carbon dioxide emissions change over short and long periods of time than just examining the emissions themselves (J.O. Ramsay and B.W. Silverman, 2005). Before that, Goreau (1990) briefly suggested studying the changes in CO2 emissions and CO2 in the atmosphere over time by applying mathematical equations known as differential equations (Tsokos C P and Xu Y., 2009). Besides, it does not seem reasonable that a larger population would produce lower CO2 emissions. Therefore, we may pay more attention to the differential data instead of the totals.

Simple linear regression by differential data

When it comes to the differential data, we hope to understand more about the change of CO2 emissions in the time domain. To figure out which factors are causing changes in carbon dioxide emissions at a specific time, we must analyze historical time series data on carbon dioxide emissions for each attributable variable. So, we operated differential transformations on the above data and then gained the differential data for the following regression analysis (Kafle R C, 2014).

Therefore, there are five simple linear regressions that we performed in Figure 7, showing quite random patterns. For example, the R-square of the linear regression between differential data of GDP and CO2 emissions is only 0.16, while that of renewable energy is 0.33. Therefore, carbon dioxide emissions have no relation between the five independent variables when we used the differential data.

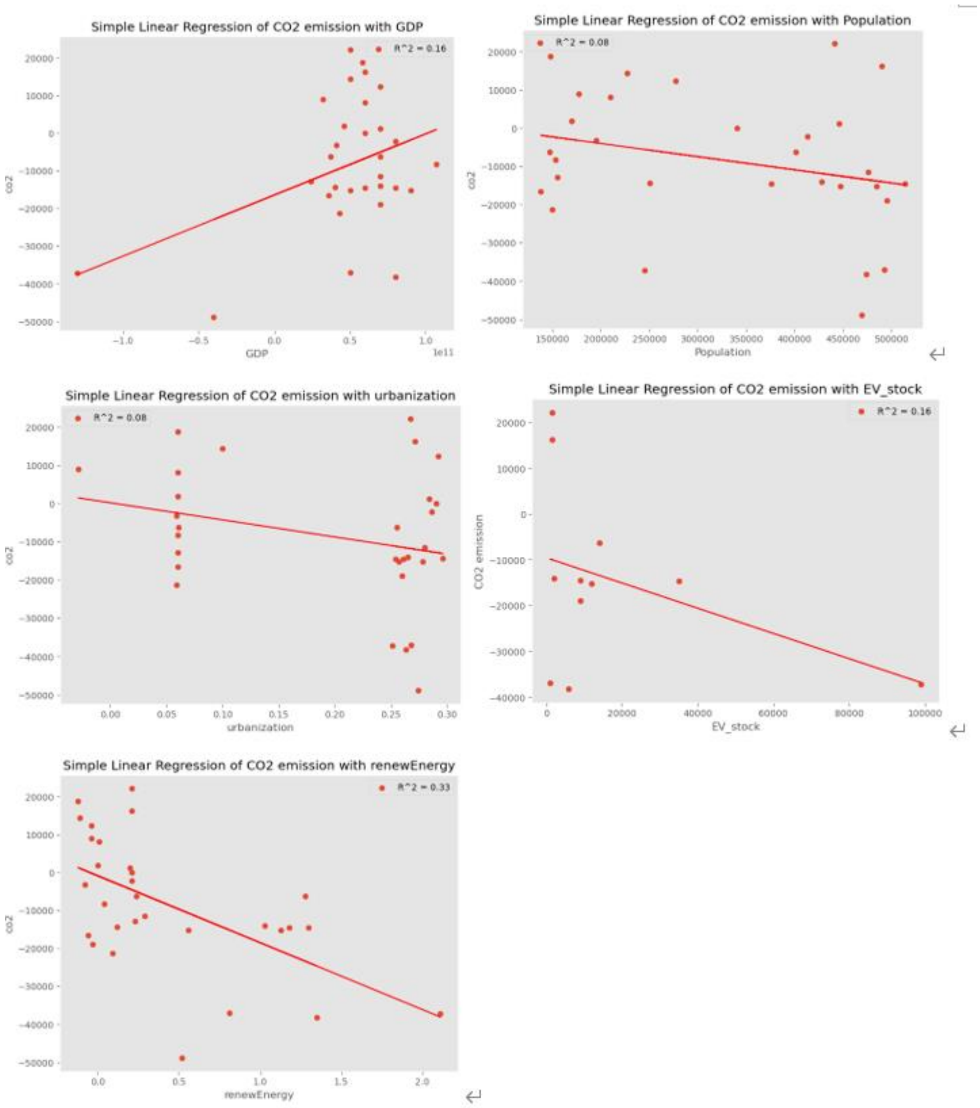

Figure 7. Simple linear regression of CO2 with the above independent variables by differential data

In conclusion, we could basically know that we would not be able to fit the differential forms of the above independent variables and that of the CO2 emissions data only by simple linear regressions.

Cluster

Since different regions have their own characteristics and different economic structures, we will perform a cluster analysis to identify carbon intensity, depending on how well they perform. A more detailed dataset, where GDP and CO2 emissions are divided by London boroughs in 2019, will be analyzed with the K-means clustering method.

Firstly, I need to standardize the GDP and CO2 variables to make sure all variables contribute equally to the K-means clustering process. In order to identify the ideal number of groups, then I computed the average silhouette width (ASW) which is a measure of clustering result. The ASW is equal to 0.685 (figure 8) when we classified data into 3 groups, indicating that the groups may be reasonable and with a suitable structure. If the number of groups move to 4, the ASW would dramatically decrease to 0.5, indicating a sign of a weak model structure.

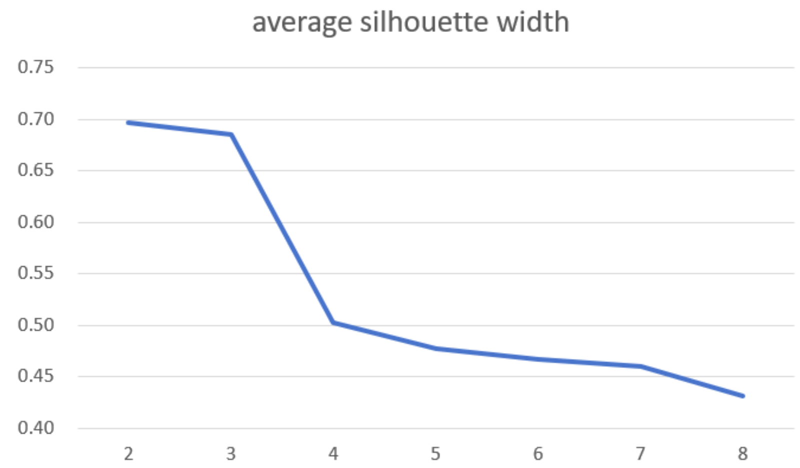

Figure 8. Average silhouette width

I finally ran the K-means clustering method and gained the cluster result in Figure 9.

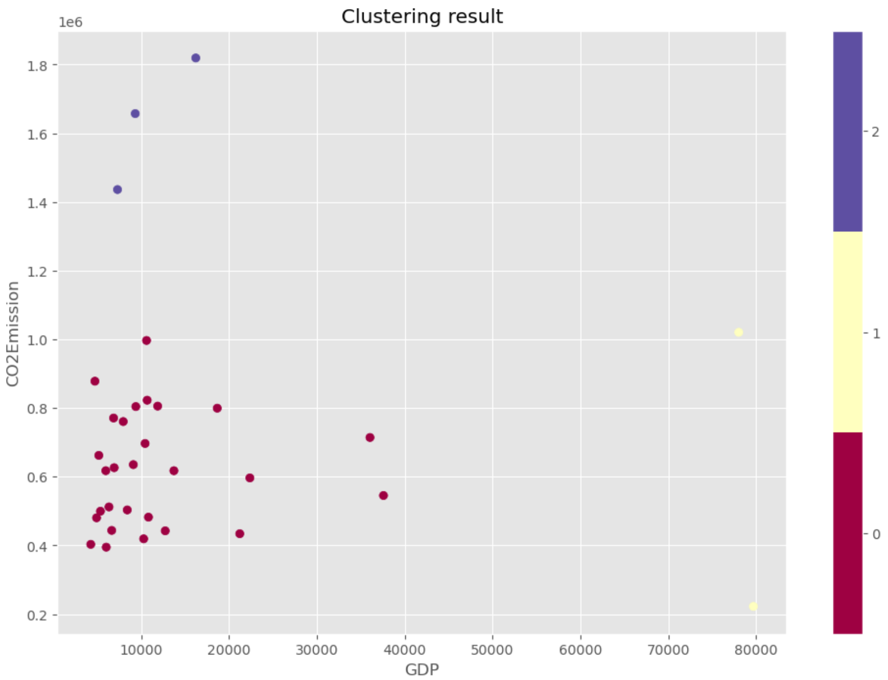

Figure 9. Cluster result of different carbon-efficient economy areas

Two boroughs are in Cluster 1, which displays the high carbon-efficient economy areas, while there are three in Cluster 2 that illustrate higher carbon intensity.

In this case, we can figure out Westminster and the city of London with higher carbon-efficient economy, because they produced around 80000 million pounds in 2019 and generated little CO_2. Westminster is also expected to continue to retain its carbon-efficient economy, largely due to its status as a center of commercial, business and tourist activity. Meanwhile, with a current GDP of £80,000 million, the City of London is expected to maintain its high carbon-efficient economy, as a financial powerhouse.

Besides, Bexley, Enfield and Hillingdon would be clustered into group 2 because of their higher CO_2 emissions and lower GDP attribution. For example, Hillingdon produced over 1819 million tonnes of CO_2 in 2019 which is 2.5 times higher than the mean of London data by boroughs (713 million) and only created 16,203 million pounds in this year. These high-emission areas are partly due to the presence of more industries, high-emission vehicles or even airports.

Some measurements such as the Ultra Low Emission Zone (ULEZ), can be applied to improve this problem. This solution has recently been extended to Greater London (2023), with additional charges imposed on vehicles that do not meet emissions standards. Although ULEZ is improving London's air quality, benefiting public health and helping to combat climate change, Hillingdon and Bexley councils have announced their resolve to work with other outer London boroughs to resist its implementation.

Conclusion

Overall, five independent variables (BEVs, GDP, urbanization, REC and population) show higher correlations with the CO_2 emissions in the UK, while we have ignored the changes in the time domain. However, the models created by differential data cannot directly apply to predict the CO_2 emissions because of low accuracy.

Finally, the cluster analysis has divided into three groups 33 borough areas in London, a good structure that could give indications about the economic efficiency and environmental costs. The policymaker could pay more attention to those lower efficiency areas and learn from the higher performance places.

In the future, we could account for other important economic indicators which can also influence CO_2 emissions, while the longer time series dataset would be gained to support better analysis. For instance, the BEVs data has only a 10-year period, because the electric car is the emerging item. Besides, although the clustering result has significant implications for policymakers, the research areas should expand rather than only include London. Finally, we may use more complex regression models to gain better results.

References

Dr Amina Syed, the decoupling of economic growth from carbon emissions: UK evidence, 2019

Hannesson R. CO2 intensity and GDP per capita[J]. International Journal of Energy Sector Management, 2020, 14(2): 372-388.

J.O. Ramsay and B.W. Silverman, Functional Data Analysis, 2nd ed. Springer, New York, 2005.

Kafle R C. Trend Analysis and Modeling of Health and Environmental Data: Joinpoint and Functional Approach[M]. University of South Florida, 2014.

Ram C. Kafle, Keshav P. Pokhrel, Netra Khanal & Chris P. Tsokos (2019) Differential equation model of carbon dioxide emissions using functional linear regression, Journal of Applied Statistics, 46:7, 1246-1259, DOI: 10.1080/02664763.2018.1542667

Sun D, Kyere F, Sampene A K, et al. An investigation on the role of electric vehicles in alleviating environmental pollution: evidence from five leading economies[J]. Environmental Science and Pollution Research, 2023, 30(7): 18244-18259.

Thomas J. Goreau, Balancing atmospheric carbon dioxide., Ambio 19 (5) (1990) 230–236.

Tsokos C P, Xu Y. Modeling carbon dioxide emissions with a system of differential equations[J]. Nonlinear Analysis: Theory, Methods & Applications, 2009, 71(12): e1182-e1197.

Yandle, B., Vijayaraghavan, M. and Bhattarai, M., 2004. Environmental Kuznets Curves: A Review of Findings, Methods, and Policy Implications. A Primer, PERC Research Study 02-01

Zafar MW, Shahbaz M, Sinha A, Sengupta T, Qin Q (2020) How renewable energy consumption contribute to environmental quality? The role of education in OECD countries. J Clean Prod 268:122149. https://doi.org/10.1016/J.JCLEPRO.2020.122149

Word Count (excluding list of references):	*2094*