CHEMICAL ENGINEERING TRANSACTIONS

VOL. 106, 2023

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Jeng Shiun Lim, Nor Alafiza Yunus, Peck Loo Kiew, Hon Huin Chin

ISBN 979-12-81206-05-2; **ISSN** 2283-9216

DOI: 10.3303/CET23106031

Pearson Correlation Analysis Between Carbon Dioxide Emissions and Socioeconomic Factors Across Nations' Income Groups

Yuting Qin[a], Min Yee Chin[a], Zheng Xuan Hoy[a], Chew Tin Lee[b], Musharib Khan[c], Kok Sin Woon[a,*]

[a]School of Energy and Chemical Engineering, Xiamen University Malaysia, Jalan Sunsuria, Bandar Sunsuria, 43900 Sepang, Selangor Darul Ehsan, Malaysia

[b]School of Chemical and Energy Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, 81300 Johor Bahru, Johor Darul Takzim, Malaysia.

[c]Institute of Environmental Sciences and Engineering (IESE), School of Civil and Environmental Engineering (SCEE), National University of Sciences and Technology (NUST), Sector H-12, Islamabad, Pakistan

koksin.woon@xmu.edu.my

Excessive carbon dioxide (CO_2) emissions are driving unprecedented climate changes, necessitating a comprehensive understanding of the underlying socioeconomic factors to develop effective policies for limiting global CO_2 concentrations. Current research lacks a broad global perspective that classifies countries based on income levels, limiting our ability to identify specific challenges and opportunities within different economies. This study aims to fill this gap by investigating the correlation between CO_2 emissions and socioeconomic factors in ten (10) high-income and lower-middle-income countries. Through Pearson correlation analysis, the study examines the relationship between CO_2 emissions and key indicators such as gross domestic product, population, energy consumption, renewable energy usage, and energy intensity from 1991 to 2021. Among the ten countries studied, the socioeconomic factors of high-income countries are more negatively correlated to carbon emissions, while lower-middle-income countries are more positively correlated. Both income groups have a positive correlation between energy consumption and carbon emissions. This research contributes to evidence-based policymaking and promotes sustainable development practices, offering valuable guidance for policymakers worldwide to combat climate change and reduce carbon emissions across diverse income groups.

1. Introduction

Carbon dioxide (CO_2) and other greenhouse gas (GHG) emissions are major contributors to climate change, posing a significant global threat. The IPCC 6th Assessment Report stated that human activities have caused an average global temperature increase of 1.1 °C since the pre-industrial era, starting in 1750. Projections indicate a potential exceedance of 1.5 °C in the short term, highlighting the urgency of addressing global warming. Understanding the factors driving CO_2 emissions is crucial for developing effective policies to mitigate climate change. According to Woon et al. (2023), a country's CO_2 emissions are primarily influenced by economic development, population dynamics, and energy utilisation.

Previous research has explored the relationship between various factors and CO_2 emissions. Salari et al. (2021) established a strong correlation between CO_2 emissions, Gross Domestic Product (GDP), and energy consumption in the United States, illustrating a mutually influential relationship between CO_2 emissions and energy usage while ensuring sustainable economic growth. Cai et al. (2018) utilised the autoregressive distributive lag bound test to examine the relationship between renewable energy usage, GDP, and CO_2 emissions in advanced economies, i.e., G7 countries. Aviso et al. (2021) detected the cities' energy utilisation and GHG emissions patterns with rough set-based machine learning. Different from past studies that focused on limited countries or regions, this research provides a novel approach by globally analysing carbon emissions in different income groups. By considering variations in economic structures, energy consumption patterns, and

Paper Received: 1 June 2023; Revised: 25 July 2023; Accepted: 10 October 2023
Please cite this article as: Qin Y., Chin M.Y., Hoy Z.X., Lee C.T., Khan M., Woon K.S., 2023, Pearson Correlation Analysis Between Carbon Dioxide Emissions and Socioeconomic Factors Across Nations' Income Groups, Chemical Engineering Transactions, 106, 181-186 DOI:10.3303/CET23106031

policy priorities, it overcomes the territory limitations of previous research, allowing for a more comprehensive understanding of emission patterns and underlying factors to uncover unique challenges and opportunities specific to each income group.

To address this research gap, this study investigates the correlation between CO_2 emissions and driving factors in the top five high-emitting countries from high-income and lower-middle-income countries. The inclusion of the lower-middle-income group, rather than the low-income group, is justified by their significantly larger CO_2 emissions volume from the former (top five countries accounting for 12.37 % of total emissions) compared to the latter (top five countries accounting for only 0.25 % of total emissions). Lower-middle-income countries have a larger population share (43 % of the world's population) compared to low-income countries (9 % of the world's population) and exhibit faster economic growth (5.6 % annual GDP growth compared to 2.6 % for low-income countries). Some lower-middle-income countries may have achieved some level of development success, offering increased opportunities and potential in carbon emissions research.

The Pearson correlation coefficient test is employed to determine the direction and strength of the correlation between two variables (Asuero et al., 2006). This method has been widely used in similar studies, evaluating the linear association between continuous variables. Li et al. (2022) applied Pearson correlation coefficient test to determine the primary socioeconomic factors that advocate CO_2 emissions and the urban compactness degree of the factors in Chinese cities. Tian et al. (2022) utilised Pearson correlation analysis to analyse the relationship between socioeconomic factors and e-waste generation before forecasting future trends. Hoy et al. (2022) used Pearson correlation analysis to show that each solid waste physical composition exhibits collinearity with different socioeconomic indicators.

A clear understanding of the socioeconomic factors contributing to increasing CO_2 concentrations is essential for developing policies to limit, or at least stabilise, global carbon concentrations. This study aims to analyse the correlation between CO_2 emissions and five socioeconomic factors – population, GDP, energy consumption, renewable energy, and energy intensity level – across different nations' income levels. By focusing on high-income and lower-middle-income countries, this research provides valuable insights into the driving forces behind carbon emissions in each income group. The findings facilitate the development of targeted climate change mitigation policies to address the specific challenges and opportunities faced by different income groups, eventually supporting worldwide initiatives to mitigate climate change and advance sustainable development.

2. Methodology

Figure 1 depicts the framework of this research. The initial phase involves collecting data on the annual carbon emissions of ten countries and five drivers from 1991 to 2021. In phase 2, data processing has been conducted to facilitate subsequent correlation testing. In Phase 3, the annual carbon emissions of each country are subjected to a Pearson's correlation test with each of the five drivers to assess the linear correlation.

2.1 Phase 1: Data collection

The study investigates the correlation between carbon emissions and socioeconomic factors across two income levels, encompassing ten countries over 31 years, from 1991 to 2021. Ten countries have been chosen to represent the top five carbon emitters in 2019, including high-income countries: the United States (USA), Japan, Germany, Saudi Arabia, and South Korea, and lower-middle-income countries: India, Iran, Indonesia, Vietnam, and Egypt. Data on carbon emissions that were obtained from Our World in Data (Ritchie et al., 2020) reflects the total emissions across all sectors. The data source reveals that, over the years, most countries experienced an increase in carbon emissions, except for the United States and Germany displayed a decreasing trend, while Japan presented a relatively stable trend. The chosen socioeconomic factors for analysis are based on established factors supported by previous literature, including population (Sarkodie et al., 2020), GDP (Salari et al., 2021), energy consumption (Salari et al., 2021), renewable energy consumption (Cai et al., 2018), and energy intensity level of primary energy (Sarkodie et al., 2020). These factors' definitions are summarised in Table 1, while their data are sourced from the World Development Indicators database (World Bank, 2023).

2.2 Phase 2: Data processing

Data preparation before the correlation test involves two essential steps: data cleaning and linearity checking. Data cleaning involves removing outliers and missing data to ensure accurate correlation analysis. Scatterplots are used to visually identify data points that deviate significantly from linearity. These outliers, which greatly differ from the overall pattern, are eliminated to minimise their impact on the overall trend. Linearity checking ensures a linear relationship between the correlated variables. Scatter plots reveal an inflection point around the year 2000 for high-income countries' correlation between carbon emissions and renewable energy use. Analysing data before this turning point may yield inaccurate results. To align with global development patterns, the correlation was analysed using data after the observed inflection point.

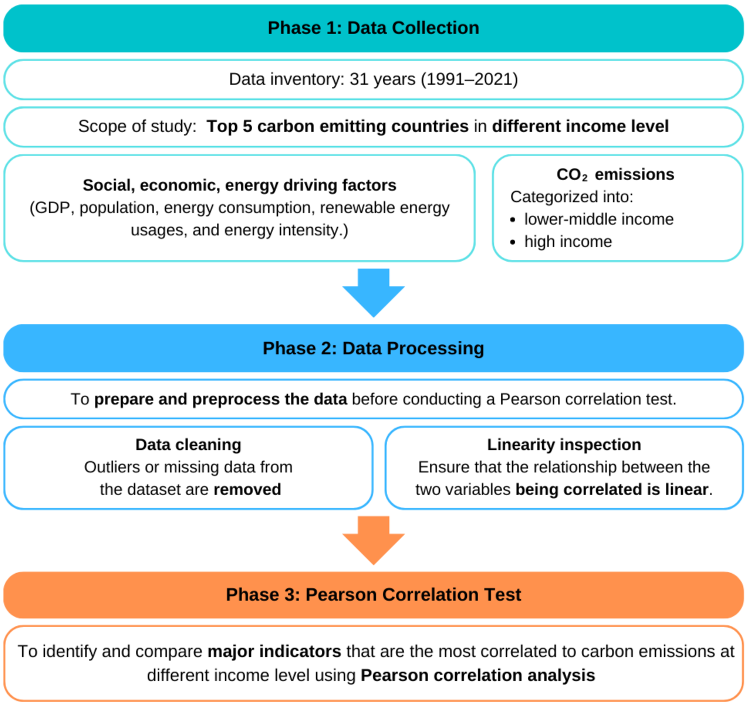


Figure 1: Overall methodology flow of the research

Table 1: Data source of driving factors for correlation analysis

Socioeconomic factors	Unit	Definition from World Bank
Population	(Number of individuals)	The total population includes all citizens regardless of legal status or citizenship, according to the de facto definition. The values provided are mid-year projections.
GDP	(Current USD)	The total value of goods and services produced by the producer in an economy, including taxes paid by producers but excluding subsidies received, in current U.S. dollars.
Energy consumption	(kg of oil equivalent per capita)	Energy consumption encompasses the initial utilisation of primary energy before its conversion into various end-use fuels. It equals to indigenous production with imports and stock changes, but excluding the export fuels and the energy used for ships and airplanes involved in international travel.
Renewable energy consumption	(% of total final energy consumption)	Renewable energy consumption refers to the proportion of renewable energy within the overall final energy consumption.
Energy intensity level of primary energy	(MJ/USD 2017 GDP at purchasing power parity)	Energy supply to GDP ratio measured at purchasing power parity. Energy intensity shows energy used per economic unit. A lower ratio indicates less energy per unit output.

2.3 Phase 3: Pearson correlation test

Pearson correlation is applied in this study using the correlation analysis tool provided in Microsoft Excel. It measures the strength and direction of the linear relationship between the selected socioeconomic factor and the nation's annual carbon emissions. The correlation coefficient ranges between -1 (perfectly negative correlation) and 1 (perfectly positive correlation), with 0 denoting no correlation. The closer the coefficient is to -1 or 1, the stronger the correlation. The formula for the Pearson correlation test is shown as Eq(1) (Asuero et al., 2006).

$$r_{xy} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}} \quad (1)$$

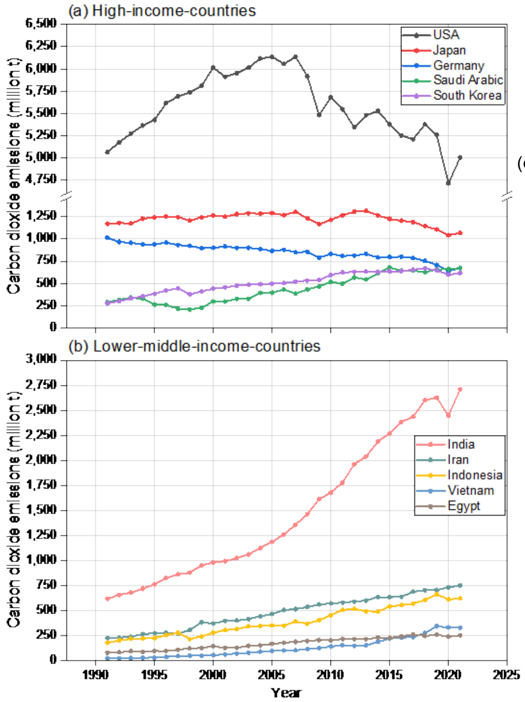
where,

r_{xy} – correlation coefficient of the linear relationship between variables x and y

x_i – values of x-variable in the data
$\bar{x}$ – mean value of x-variable
y_i – values of y-variable in the data
$\bar{y}$ – mean value of x-variable

3. Results and discussion

The results of the correlation analysis between carbon emissions and socioeconomic factors are depicted in Figure 2(c), with the intensity of colour indicating the strength of the correlation. The blue colour indicates a positive correlation, while the red colour implies a negative correlation. A correlation coefficient greater than 0.8 generally indicates a strong correlation between two factors (Frost, 2019).



(c) Correlation coefficient test results

	USA	JPN	DEU	SAU	KOR	IND	IRN	IDN	VNM	EGY
Pop	-0.935	0.458	-0.458	0.948	0.976	0.968	0.990	**0.987**	0.934	**0.979**
GDP	-0.872	0.106	-0.879	0.946	0.939	0.991	0.752	0.971	0.980	0.908
EC	**0.989**	0.275	0.913	**0.949**	**0.992**	**0.997**	**0.990**	0.943	**0.996**	0.943
RE	-0.934	-0.542	**-0.950**	-0.005	0.795	-0.943	-0.151	-0.983	-0.933	-0.967
EI	0.927	**0.561**	0.945	-0.777	-0.877	-0.940	0.870	-0.952	-0.583	-0.111
	High-income					Lower-middle-income				

+1 positive correlation ... -1 negative correlation

Figure 2: (a) CO_2 emissions trend for high-income countries, (b) CO_2 emissions trend for lower-middle-income countries, (c) socioeconomic correlation coefficient of the studied countries under high-income and lower-middle-income levels. The bold values indicate the strongest correlation between the particular socioeconomic factor and the carbon emissions within the country. Note that USA: the United States; JPN: Japan; DEU: Germany; SAU: Saudi Arabia; KOR: South Korea; IND: India; IRN: Iran; IDN: Indonesia; VNM: Vietnam; EGY: Egypt; Pop: population; GDP: gross domestic product; EC: Energy consumption; RE: The share of renewable energy consumption; EI: Energy intensity level of primary energy.

3.1 High-income-countries

In the high-income group, the USA and Germany, which have a decreasing trend, as illustrated in Figure 2(a), exhibit negative correlations with population, GDP, and renewable energy, in contrast to Saudi Arabia and South Korea. Japan shows weak correlations with all five drivers tested, with correlations below 0.5 in absolute terms, except for a correlation of -0.542 with renewable energy sources and 0.561 with energy intensity. As shown in Figure 2(a), Japan's CO_2 emissions have remained relatively stable between 1 Gt and 1.3 Gt from 1990 to 2021 (World Bank, 2023). Unlike other countries, Japan's emissions trajectory does not show pronounced variation, suggesting a consistent pattern over the analysed period. This limited variability in Japan's carbon emission trends is the primary factor contributing to the low magnitude of the linear correlation coefficient. The emission trend remained unaffected despite changes in socioeconomic factors. Japan's relatively stable carbon emission trend is due to the dominant role of nuclear energy and efforts to improve energy efficiency. Historically relying on nuclear power, Japan has consistently maintained low emissions. Despite the Fukushima nuclear disaster, energy efficiency measures and technologies were implemented to conserve energy and sustain low emissions. Its industrial structure and strict regulations further contribute to the low carbon footprint.

Among the five high-income countries, four demonstrate a negative correlation between carbon emissions and renewable energy sources. Germany exhibits the highest correlation coefficient of -0.950, followed by the United States (-0.934), Japan (-0.542), Saudi Arabia (-0.005), and South Korea (0.795). With the decreasing trend in CO_2 emissions, the results show that the United States, Germany, and Japan have increasingly invested in renewable energy. These findings highlight the overall effectiveness of renewable energy adoption in mitigating carbon emissions across most countries analysed. South Korea exhibits a positive correlation coefficient of 0.795, indicating its emphasis on renewable energy but with limited impact on mitigating CO_2 emissions. The escalating energy demand in South Korea, driven by its status as a developing nation, is outpacing the growth rate of new energy sources. Consequently, despite efforts to develop renewable energy, the continued consumption of conventional energy sources leads to a persistent increase in CO_2 emissions.

Germany stands out as a noteworthy example where the utilisation of renewable energy plays a pivotal role in driving carbon emissions reduction efforts. The German government has steadfastly committed to mitigating carbon emissions and actively promotes developing and implementing renewable energy technologies. Germany has achieved a substantial share of new energy installations, a significant portion of flexible regulated power sources, and a relatively stable system load. As of 2019, new energy generation accounted for over 30 % of the country's total electricity consumption (Yuan et al., 2020). The high correlation coefficient value reflects Germany's notable achievements in energy transition and sustainable development.

Saudi Arabia shows a limited and insignificant correlation between CO_2 emissions and renewable energy usage. This disparity can be attributed to the comparatively slower progress made by Saudi Arabia in developing alternative energy sources, as they remain heavily reliant on traditional oil and gas energy resources. Diversifying energy sources and pursuing sustainable development remain crucial for Saudi Arabia. Increased investment in renewable energy sources and technological innovation is necessary to reduce their dependence on fossil fuels and decrease carbon footprints (Koide et al., 2021).

3.2 Lower-middle-income countries

The association between population, economic indicators, and carbon emissions in lower-middle-income countries demonstrates a more substantial relationship than in high-income countries. Within the lower-middle-income group, population size exhibits a strong positive correlation with carbon emissions, with correlation coefficients exceeding 0.9 for all countries. Population size is a significant driver of carbon emissions in Indonesia and Egypt, with correlation coefficients of 0.987 and 0.979, respectively. GDP displays a robust positive correlation within the lower-middle-income group, with all countries exhibiting correlation coefficients above 0.9, except for Iran (0.752). These findings align with the study conducted by Sarkodie et al. (2020), which highlights the importance of population dynamics and economic development in shaping carbon emissions, particularly in lower-middle-income countries. The reliance on traditional fossil-based energy sources, coupled with growing economies and populations, leads to increased energy demand and inefficient energy consumption and production, resulting in higher carbon emissions. Lower-middle-income countries must implement measures to address this issue, including promoting more efficient energy use, improving energy infrastructure, promoting clean energy, and strengthening energy management.

Examining the five countries in the lower-middle-income group, a simultaneous decline in the utilisation of renewable energy sources is observed alongside a persistent upward trend in carbon emissions, as presented in Figure 2(b). The negative correlation between the share of renewable energy use and carbon emissions in this income group suggests a high reliance on non-renewable energy sources. This pattern arises from the fact that while the overall quantity of energy sources is increasing, fossil energy consumption is rising faster, leading to a decline in the share of renewable energy use. Lower-middle-income countries are currently in the early stages of development, and most have not yet undergone an energy transition. The countries heavily rely on fossil fuels to sustain their development. Achieving a successful energy transition requires substantial investment and technological support, and these countries face various challenges in realising this goal, such as limited access to technology and funding.

4. Conclusions

This study employed Pearson correlation analysis to examine the relationship between socioeconomic factors and carbon emissions. In high-income countries, specifically the United States, Japan, and Germany, carbon emissions significantly negatively correlated with population, GDP, and the share of renewable energy usage, while positively correlating with energy consumption and intensity. In lower-middle-income countries, carbon emissions were positively correlated with population, GDP, and energy consumption but negatively correlated with the proportion of renewable energy usage and energy intensity. These results indicate that energy-related factors play a crucial role in carbon emissions for high-income countries, while population and economic factors drive emissions in lower-middle-income countries. Energy consumption emerged as the most influential factor across countries, with a strong positive correlation coefficient exceeding 0.9, except for Japan. Policymakers can develop effective strategies for managing carbon emissions by analysing and comparing the highest correlation factors. High-income countries should prioritise renewable energy initiatives and energy efficiency advancements, taking a leadership role in advancing technology and fostering innovation. On the other hand, lower-middle-income countries should strive for a balance between economic development and energy consumption. They need comprehensive strategies for sustainable development and expedited transition to more sustainable energy sources. This study enhances understanding of carbon emissions dynamics and guides policy decisions towards sustainable development. Future research can utilise the outcomes of correlation analysis to project future carbon emissions by incorporating highly correlated factors as inputs into the forecasting process for each specific country.

Acknowledgements

The authors would like to thank for the financial support from Xiamen University Malaysia Research Fund (XMUMRF/2019-C4/IENG/0022 and XMUMRF/2023-C11/IENG/0057).

References

Arias P., Bustamante M., Elgizouli I., Flato G., Howden G., Méndez C., Pereira J., PichsMadruga R., Rose S.K., Saheb Y., Sánchez R., Ürge-Vorsatz D., Xiao C., Yassaa N., 2023, Synthesis Report of the IPCC Sixth Assessment Report (AR6), Intergovernmental Panel on Climate Change <report.ipcc.ch/ar6syr/pdf/IPCC_AR6_SYR_LongerReport.pdf> accessed 11.03.2023.

Asuero A.G., Sayago A., González A.G., 2006, The correlation coefficient: An overview, Critical Reviews in Analytical Chemistry, 36(1), 41–59.

Aviso K.B., Capili M.J., Chin H.H., Fan Y.V., Klemeš J.J., Tan R.R., 2021, Detecting Patterns in Energy Use and Greenhouse Gas Emissions of Cities Using Machine Learning, Chemical Engineering Transactions, 88, 403-408.

Cai Y., Sam C.Y., Chang T., 2018, Nexus between clean energy consumption, economic growth and CO2 emissions, Journal of Cleaner Production, 182, 1001–1011.

Frost J., 2019, Introduction to Statistics: An Intuitive Guide for Analysing Data and Unlocking Discoveries, Statistics by Jim Publishing, Pennsylvania, USA, 95–102.

Hoy Z.X., Woon K.S. Chin W.C., Hashim H., Fan V.Y., 2022, Forecasting heterogeneous municipal solid waste generation via Bayesian-optimised neural network with ensemble learning for improved generalisation, Computers & Chemical Engineering, 166, 107946.

Jorgenson A.K., Fiske S., Hubacek K., Li J., McGovern T., Rick T., Schor J.B., Solecki W., York R., Zycherman A., 2019, Social science perspectives on drivers of and responses to global climate change, Wiley Interdisciplinary Reviews: Climate Change, 10(1), e554.

Li Z., Wang F., Kang T., Wang C., Chen X., Miao Z., Zhang L., Ye Y., Zhang H., 2022, Exploring differentiated impacts of socioeconomic factors and urban forms on city-level CO2 emissions in China: Spatial heterogeneity and varying importance levels, Sustainable Cities and Society, 84, 104028.

Koide R., Lettenmeier M., Akenji L., Toivio V., Amellina A., Khodke A., Watabe A., Kojima S., 2021, Lifestyle carbon footprints and changes in lifestyles to limit global warming to 1.5 °C, and ways forward for related research, Sustainability Science, 16(6), 2087-2099.

Ritchie H., Roser M., Rosado P., 2020, CO_2 and greenhouse gas emissions, Our World in Data <ourworldindata.org/co2-and-greenhouse-gas-emissions> accessed 24.03.2023.

Salari M., Javid R.J., Noghanibehambari H., 2021, The nexus between CO2 emissions, energy consumption, and economic growth in the U.S., Economic Analysis and Policy, 69, 182–194.

Sarkodie S.A., Owusu P.A., Leirvik T., 2020, Global effect of urban sprawl, industrialisation, trade and economic development on carbon dioxide emissions, Environmental Research Letters, 15(3), 034049.

The World Bank, 2023, World Development Indicators, The World Bank Group <databank.worldbank.org/source/world-development-indicators> accessed 24.03.2023.

Tian R., Hoy Z.X., Woon K.S., Hanafiah M.M., Liew P.Y., 2022, Socioeconomic Correlation Analysis of E-waste and Prediction of E-waste Generation via Back-propagation Neural Network, Chemical Engineering Transactions, 97, 397–402.

Woon K.S., Phuang Z.X., Taler J., Varbanov P.S., Chong C.T., Klemeš J.J., Lee C.T., 2022, Recent Advances in Urban Green Energy Development Towards Carbon Neutrality, Energy, 126502.

Yuan W., Wang C., Ye X., Li Q., Xu L., Li Z., Shi Z., 2020, Practice and experience in achieving a high proportion of renewable energy in Germany, E3S Web of Conferences, 185, 01001.