Data Mining Analysis of Changes in Energy Consumption by Transportation Structure Adjustment: An Index Decomposition Approach

Jian Jiang[1], Yun Jing[1]*, Qing Liu[2], Yu Tian[2], Jing Wang[2]

[1]School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

[2]Communication and Signal Research Institute, China Academy of Railway Sciences Group Co., Ltd, Beijing, China

*Corresponding Author.

Abstract

For the backdrop of rapid economic growth in China, the energy consumption it generates is also a focus of attention. The data mining analysis of this amount in transportation production department is very crucial and transport has always accounted for a considerable proportion of total energy consumption. The transportation structure adjustment policy formulated by the country is aimed at this point, hoping to reduce the total energy consumption through benign adjustments to transportation methods. This article applies decomposition method for data mining to decompose the changes in energy consumption in China's transportation sector, so as to analyze the impact of transportation structure adjustment policies on it. This method establishes the relationship between total amount and factors through multiplication identities, and then analyzes the degree of influence of factors on total amount changes by combining the total amount and changes in various factors over a time period. From the calculation results, it can be seen that the transportation structure adjustment policy has played an important role in reducing energy consumption in the transportation sector, and the country has made very correct decisions to continue promoting the implementation of this policy in the coming years. In these analyses, data mining methods have identified key supporting factors in years of data, analyzed the severity of the impact, and played an important role in the article.

Keywords: Index decomposition, energy consumption, transportation structure adjustment, transport intensity, data mining

1. Introduction

The fast development of the Chinese economy in recent decades has attracted the attention of the world and has become the second largest economy on the earth since the year 2010. Meanwhile, the energy consumption and carbon emissions in China are gradually among the top in the world, with the energy consumption by the unit of GDP at a relatively high level. In September 2020, China clearly proposed the goals of "peaking carbon emissions by 2030" and "carbon neutrality" by 2060. As a result, both the CO_2 emissions and energy consumption per unit of GDP in China has decreased by more than 40% compared to 2012, and the total energy consumption was reasonably controlled, resulting in an improvement in clean, low-carbon, and efficient levels. In 2022, the share of non-fossil fuels increased by 7.8 percentage compared to 2012 in China, and the energy consumption structure gradually optimized, and energy utilization efficiency steadily improved.

The transportation sector is very important and crucial for energy consumption, and China's current achievements in energy utilization and CO_2 emission reduction are partly attributed to the contribution of the transportation sector. In September 2018, after the State Council released the "Three Year Action Plan for Promoting

Transportation Structure Adjustment (2018-2020)", the transportation department vigorously carried out transportation structure adjustment, taking the "road to railway" as a breakthrough point, and striving to promote the formation of the most suitable transportation method to complete the transportation of goods.

Data Mining (DM) refers to the process of searching for information hidden within a large amount of data through algorithms. Data mining is usually related to computer science and is achieved through various methods such as statistics, online analytical processing, intelligence retrieval, machine learning, expert systems (relying on past empirical rules), and pattern recognition. Data mining is an interdisciplinary field that involves various fields such as database technology, artificial intelligence, machine learning, pattern recognition, high-performance computing, knowledge engineering, neural networks, information retrieval, and information visualization. The standard process of data mining includes six stages: business understanding, data understanding, data preparation, data modeling, model evaluation, and deployment. By applying this technology, key factors affecting energy consumption in transportation systems can be identified and analyzed.

Among many influencing factors, the adjustment of transportation structure will transfer a portion of the turnover of highway transportation to railways and waterways, thereby reducing the overall energy consumption density. The result of transportation structure adjustment for energy conservation can be analyzed by analyzing the changes in comprehensive energy consumption density and their impact on energy consumption changes. Consequently, there is necessary to analyze the effect of transportation structure adjustment on energy consumption quantitatively and we need to conduct quantitative research in this situation to obtain actual numerical values to analyze the effect of every specific influencing factor and propose corresponding policy improvement suggestions.

This article analyzes the dominant factors behind the change amounts in energy consumption in China's transportation industry in recent years, as well as the impact of national transportation structure adjustment policies on energy consumption changes. Therefore, we have made appropriate improvements to the methodology of factor decomposition analysis to make it more suitable for the objective characteristics of various transportation modes. We will provide a detailed explanation of the details of this aspect in the next section, and the decomposition results, analysis, and policy recommendations will be provided at the end of the article.

As mentioned earlier, the energy consumption of the transportation sector is quantitatively influenced by the energy intensity of transportation vehicles, traffic turnover, and gross domestic product as well. Then there must be some mathematical relationship between the change amount in energy consumption and the change amounts in these influencing factors. In other words, the variation of each influencing factor together constitutes and determines the total variation. The decomposition analysis method focuses on the split of changes in the total amount, and decomposes the changes in the total amount into the total of the effects brought by the changes in various factors according to certain decomposition rules. Therefore, the changes in transportation energy consumption will be split into multiple predefined factors through decomposition analysis [1]. Then, in this article, the amount of energy consumption changes in the transportation sector are decomposed into three factors. Therefore, the changes in transportation energy consumption can also be quantitatively displayed by the changes in factors, and the quantitative analysis results can reflect the contribution of each factor. Decision makers can use this as a basis to identify key factors that contribute to the transportation energy consumption, and make targeted improvements to balance energy conservation and economic development.

The exponential decomposition method was initially used to study the impact of changes in product structure on industrial energy demand. In these years, it is not only widely used in the industrial field, but also in the environmental field [2-4]. It mostly deals with problems where the total amount can be expressed by multiplying factors. Based on the current research reports on this method, it is not difficult to find that index decomposition analysis is widely accepted as an analytical tool for energy and environmental decision-making for a nation [5]. With the development of this method, many mathematical algorithms have been applied to this decomposition calculation. Using the Logistic Mean Division Index (LMDI) model to explore the driving forces behind changes in CO_2 emissions in transportation sector from a temporal perspective during 2000-2017[6]. Besides, the heterogeneity and spatial autocorrelation CO_2 emissions in transportation sector of 51 "Belt and Road" economic entities from the year 2000 to 2014 is examined [7].

The amount of the emission of carbon dioxide and energy consumption is very important in analyzing the relationship between economic development, transportation turnover, and environmental sustainable development. Their characteristics are also very similar, so their quantitative analysis methods are very similar and can be referenced and used for reference [8-12]. For data analysis, dynamic models are applied to estimate through means of the standard generalized method of moments and system generalized method of moments [13]. At the case of data analysis study, data analyses are designed to describe a formal mechanism based on design principles [14].

2. Methodology

This index decomposition method is applied to a value which can be expressed by some factors multiplying together. The core idea is to break down the changes in the total amount according to certain rules based on the various influencing factors of the total amount, and each value after splitting corresponds to the contribution made by the changes in each influencing factor. In this way, we can know how much impact each change in influencing factors has on the overall change, and thus identify the importance of each influencing factor.

In this paper, we want to find out how the transportation structure adjustment affects to the change of total transportation energy consumption(E). Therefore, the value to be decompose is the change of total transportation energy consumption. On the other hand, we suggest the factor "comprehensive energy intensity(I)" to describe the effect of transportation structure adjustment, because the adjustment can encourage the conversion between different transportation modes, and the comprehensive energy intensity will change due to the differences energy intensity and shares of the transportation modes. In addition, through the study of many literature, it is fully demonstrated that there is a very important positive correlation between a country's economic growth and energy consumption [15-18]. Then the factor "GDP(G)" which represents the economic growth, will also be considered as one factor of the energy consumption. In this way, the left factor is named "transport intensity of the GDP(S)", measures the comprehensive respects of all factors that are affecting the transportation volume, and will make E=I*G*S.

From the year 0 to T, the change of E, that is ΔE, will be managed to expressed by the change of I,G,T in the same period. We start from the equation E=I*G*S. In the year 0, we have $E^0 = I^0 * G^0 * S^0$, and in the year T, we have $E^T = I^T * G^T * S^T$, then the change of E from the year 0 to T can be shown as Equation 1:

$$\Delta E = E^T - E^0 = I^T * G^T * S^T - I^0 * G^0 * S^0 \tag{1}$$

The aim of this decomposition method is to find out the effects imposed on E by each factor, especially I. So this paper decompose ΔE into the 3 items $\Delta E_I, \Delta E_G, \Delta E_S$, that express the change of the factors I,G,S ($\Delta I, \Delta G, \Delta S$) contributes to ΔE respectively.

Apparently that $\Delta I = I^T - I^0, \Delta G = G^T - G^0, \Delta S = S^T - S^0$

Therefore we have Equation 2:

$$\Delta E = (\Delta I + I^0) * (\Delta G + G^0) * (\Delta S + S^0) - I^0 * G^0 * S^0 = \Delta I * \Delta G * \Delta S + \Delta I * \Delta G * S^0 + \Delta I * G^0 * \Delta S + I^0 * \Delta G *$$
$$\Delta S + \Delta I * G^0 * S^0 + I^0 * \Delta G * S^0 + I^0 * G^0 * \Delta S \tag{2}$$

According to the rules determined by the Refined Laspeyres index method, in the various expansions of the above equation, only one factor that changes is clearly attributed to the impact of this changing factor, while the two factors that change should be evenly distributed among the effects of these two changing factors. Finally, in the item "$\Delta I * \Delta G * \Delta S$", all the three factors have changed, this item should be evenly divided into three parts, attributed to the impact of each factor. By the way, this is a perfect decomposition without generating any residual terms.

Consequently, the decomposition result in this article is as Equation 3-5 shows:

$$\Delta E_I = \Delta I * G^0 * S^0 + \frac{1}{2} \Delta I * \Delta G * S^0 + \frac{1}{2} \Delta I * G^0 * \Delta S + \frac{1}{3} \Delta I * \Delta G * \Delta S \tag{3}$$

$$\Delta E_G = I^0 * \Delta G * S^0 + \frac{1}{2} \Delta I * \Delta G * S^0 + \frac{1}{2} I^0 * \Delta G * \Delta S + \frac{1}{3} \Delta I * \Delta G * \Delta S \tag{4}$$

$$\Delta E_S = I^0 * G^0 * \Delta S + \frac{1}{2} I^0 * \Delta G * \Delta S + \frac{1}{2} \Delta I * G^0 * \Delta S + \frac{1}{3} \Delta I * \Delta G * \Delta S \quad (5)$$

Besides, there is another decomposition formulation which is called the Logarithmic mean Divisia index method, and this decomposition result of this method is also a perfect decomposition without generating any residual terms. The decomposition result of this Logarithmic mean Divisia index method is shown in Equation 6-8:

$$\Delta E_I = \ln\frac{I^T}{I^0} * L(E^T, E^0) \quad (6)$$

$$\Delta E_G = \ln\frac{G^T}{G^0} * L(E^T, E^0) \quad (7)$$

$$\Delta E_S = \ln\frac{S^T}{S^0} * L(E^T, E^0) \quad (8)$$

In the expressions above, function $L(E^T, E^0) = \frac{E^T - E^0}{\ln E^T - \ln E^0}$, which represents the mean weight function defined as the logarithmic average of two positive numbers.

In this paper, we will do the decomposition in these two methods above and can also do comparison at the same time.

3. Data and Decomposition Result

In this study, we pick up the data from official statistics yearbook in China including **China Energy Statistical Yearbook, China Statistical Yearbook, China Transportation Statistical Yearbook** from the year 2000 up to 2022, and the data is only available by the year 2020. As the transportation structure adjustment started from the year 2018, we analyze the effect year by year after the 2018, and before that, we analyze in a 5-year gap. So the analysis object time slot is year 2000-2005,2005-2010, 2010-2015, 2015-2018, 2018-2019,2019-2020.

Since China does not make the statistics of the energy consumption by passenger and cargo transport separately, we have to split the total consumption of the transportation sector into passenger and cargo transport. Correspondingly, the transportation turnover volume also adopts the conversion turnover volume from passenger transposition into freight transportation. The conversion ratios are as follows: Railway and marine-1 capita*km=1t*km; road-1 capita*km=0.1t*km; air-1 capita*km=0.072t*km.

Besides, the volume of GDP adopts the volume of GDP brought by the transportation department, and it can reflect the connection between GDP and the transportation energy consumption more tightly.

Then, according to the methodology we have introduced, the decomposition result by two perfect decomposition method are listed in the Tables 1 and 2 and shown in Figure 1 and 2.

Table 1 Decomposition result(Refined Laspeyres index method)(tons of standard coal)

	ΔE_I	ΔE_G	ΔE_S	ΔE
2000~2005	-806.06	8229.37	265.69	7689
2005~2010	-4931.13	13127.34	-230.21	7966
2010~2015	4010.61	15952.34	-8554.95	11408
2015~2018	-597.64	11498.18	-5793.54	5107
2018~2019	1236.05	2252.60	-3196.65	292
2019~2020	-570.82	-917.16	-1112.02	-2600

Table 2 Decomposition result(Logarithmic mean Divisia index method)(tons of standard coal)

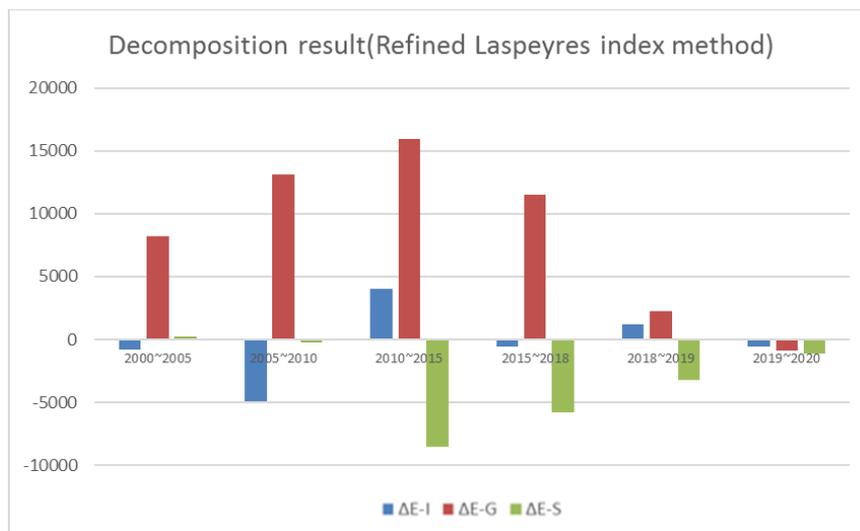|	ΔE_I	ΔE_G	ΔE_S	ΔE
2000~2005	-784.47	8214.15	259.32	7689
2005~2010	-4757.74	12947.08	-223.34	7966
2010~2015	3920.89	15761.66	-8274.55	11408
2015~2018	-592.72	11438.17	-5738.45	5107
2018~2019	1235.17	2251.01	-3194.18	292
2019~2020	-570.80	-917.16	-1112.04	-2600

Figure 1 Decomposition result(Refined Laspeyres index method)
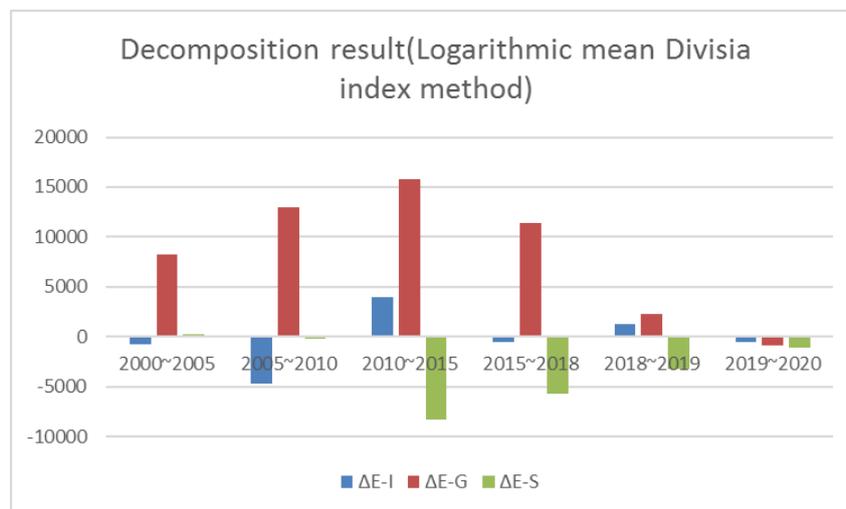
(tons of standard coal)

Figure 2 Decomposition result(Logarithmic mean Divisia index method)

(tons of standard coal)

From the two decomposition result, we can see that these two methods give almost the same result, compared with the total amount, the gap between every same item is very tiny and do not affect much to the analysis result.

4. Result Analysis and Policy Implications

Based on the decomposition results of the energy consumption increment in the transportation sector in the previous section, we can provide some interpretations. From the trend of energy consumption in the transportation sector, it can be seen that in the early 21st century, it was highly synchronized with the growth of the Chinese economy. As China began to fulfill its solemn commitment to energy conservation and emission reduction, its energy consumption intensity has greatly improved since 2010. The total energy consumption and intensity of the transportation sector have decreased to varying degrees, and since 2019, the total energy consumption of the transportation sector has decreased for the first time. Then, based on the results of factor decomposition, analyze the specific impact of each influencing factor on energy conservation.

The change in energy density is the most significant of the three factors, sometimes positive and sometimes negative, Before 2010, the rapid economic growth greatly drove energy consumption, and the transportation

energy consumption growth is also positive, but the energy density provided a negative contribution. This can be analyzed from two aspects. In terms of passenger transportation, China's railway passenger dedicated lines and high-speed railways have begun to operate, and the development of urban public transportation is also changing rapidly. The transportation capacity released by ordinary speed railways has also been invested in freight transportation, which can be considered as the reason for reducing the energy density. Between 2010 and 2015, the rapid growth of China's cargo transportation, especially the express logistics industry, has driven the explosive development of highway transportation. Although many goods have long transportation distances, due to the advantages of road transportation such as timeliness and convenience, as well as the cost advantage brought by the decrease in oil prices, road transportation has become the preferred choice for long-distance transportation, thereby increasing the contribution of energy density in this period. After 2015, the construction and operation of high-speed railways in China entered a new stage of rapid development, which brought great challenges to the road and air transportation industries. They accounted for a large share of domestic long-distance travel and also made up for the increase in energy density in freight transportation. At this period, the contribution of transportation energy density is negative. The decrease in oil prices during road transportation in 2018 was very significant, and the cost advantage of road transportation became increasingly apparent, which can be considered as the reason for the increased contribution of energy density during this stage. Although it is a market choice based on economic benefits, it is not suitable for the overall trend of energy conservation and emission reduction. This is also an important reason for the government to carry out transportation structure adjustment, using administrative means to guide the transfer of long-distance transportation to railway transportation. Therefore, in the period starting from 2019, the transportation structure adjustment achieved immediate results, and its contribution to energy density became negative, Contributed to the decrease in overall transportation energy consumption.

In addition, the contribution of GDP to transportation energy consumption is generally positive, basically exceeding the total transportation energy consumption, and the changes in transportation energy consumption are also basically synchronous. This indirectly confirms the direct relationship among GDP, economic development, and energy consumption, reflecting the positive correlation between energy consumption in the transportation sector and overall energy consumption. This trend can reflect that China has maintained a good balance between economic development and sustainable development, faithfully practicing the guiding spirit of "green mountains and clear waters are as valuable as gold and silver". The growth rate of economic development is greater than the growth rate of energy consumption, and the energy consumption and carbon emissions per unit of GDP have significantly decreased. It is a model of sustainable development, reflecting China's responsible attitude and sense of responsibility as a major country.

In terms of traffic density in GDP, this value makes the smallest contribution to traffic energy consumption and is basically negative. And it can be seen that the obvious trend is that over time, its contribution rate remains lower and lower, without exception. This indicates that China's transportation energy consumption is decreasing as it generates the same GDP in the transportation industry. At the Copenhagen Climate Conference, the Chinese government promised that by 2020, China's CO_2 emissions per unit of GDP would decrease by 40% to 45% compared to the year 2005. This also proves from the data that China is steadily making efforts to achieve this goal.

5. Conclusion and Prospects for Further Research

This article uses the exponential decomposition method to conduct a complete factor decomposition analysis on the energy consumption in China's transportation sector, so as to quantitatively analyze the contribution of the adjustment of the country's transportation structure to reducing energy consumption. The following conclusions can be drawn from the entire study:

(1). Both exponential decomposition methods can obtain perfect decomposition results, and the difference is not significant. Choosing one of the results or taking the average of both results in the application does not affect the analysis conclusion of the problem;

(2). It can be clearly seen that since China's promotion of transportation structure adjustment, the transportation energy density and its contribution to overall transportation energy consumption have had a significant decrease in impact. It can be considered that the achievements of transportation structure adjustment in sustainable development are very obvious;

(3). GDP has always been a driving factor for changes in energy consumption of transportation sector, and its contribution value is generally greater than the final increment of energy consumption. From this, it can be seen that even if transportation energy consumption increases, it creates more GDP growth. Therefore, even during the period of increasing energy consumption, its growth has contributed to economic development without unnecessary consumption. Correspondingly, the contribution of traffic density to energy consumption in GDP has always been low, reflecting that China's transport industry has been focusing on a sustainable developing model of low-carbon and low energy consumption since entering the new century, helping to save energy and reduce emissions while achieving a significant increase in transportation turnover.

Of course, there is still room for further exploration in the research of this paper:

(1). According to current statistical data, the energy consumption of passenger transportation and cargo transportation in China is not separated, which also leads to the inability to completely separate the policy effect analysis of transportation structure adjustment, which mainly focuses on freight transportation, from the overall transportation energy consumption. Passenger turnover is also converted to cargo turnover through conversion, which may cause some inaccuracy. In the future, more detailed research will be attempted through other means;

(2). Under the impact of the COVID-19, the turnover, GDP and other data of the transportation sector will change greatly from 2020 to 2022. Even with the statistical data of these years, the situation of its response problems will also be greatly affected. This is also an unavoidable problem in the quantitative analysis of time series, which needs careful consideration in data processing.

Acknowledgment

References
[1] Ang,B.W. 2004. Decomposition analysis for policymaking in energy: which is the preferred method? Energy Policy32, 1131-1139.

[2] Ang,B.W.,Zhang,F.Q.,2000. A survey of index decomposition analysis in energy and environmental studies. Energy25, 1149-1176.

[3] Ang, B.W., 2015. LMDI decomposition approach: a guide for implementation. Energy Pol. 86, 233–238.

[4] Su B, Ang BW. 2015Multiplicative decomposition of aggregate carbon intensity change using input–output analysis. Appl Energy;154:13–20.

[5] Xu XY, Ang BW. 2013. Index decomposition analysis applied to CO_2 emission studies. Ecol Econ;93:313–29.

[6] Y.X. Liu, S.Y. Yang, X.M. Liu, P.B. Guo, K.Y. Zhang.Driving forces of temporal-spatial differences in CO_2 emissions at the city level for China's transport sector.Environ Sci Pollut Res, 28 (2021), 25993-26006

[7] N. Liu, Z. Ma, J. Kang, B. Su. 2019. A multi-region multi-sector decomposition and attribution analysis of aggregate carbon intensity in China from 2000 to 2015 Energy Pol, 129, 410-421

[8] J. Wang, T. Zhao, X. Xu, X. Zhang. 2016. Exploring the changes of energy-related carbon intensity in China: an extended Divisia index decomposition. Nat Hazards, 83 (1), 501-521

[9] Faruque, Md.O., Rabby, Md.A.J., Hossain, Md.A., Islam, Md.R., Rashid, M.M.U.,Muyeen, S.M., 2022. A comparative analysis to forecast carbon dioxide emissions. Energy Rep. 8, 8046–8060.

[10] Huo, T., Xu, L., Liu, B., Cai, W., Feng, W., 2022. China's commercial building carbon emissions toward 2060: an integrated dynamic emission assessment model. Appl. Energy 325, 119828.

[11] Wu R, Dong J, Zhou L, Zhang L. 2018.Regional distribution of carbon intensity and its driving factors in China: an empirical study based on provincial data. Pol J Environ Stud;27(3).

[12] Sebri, M., Ben-Salha, O., 2014. On the causal dynamics between economic growth, renewable energy consumption, CO_2 emissions and trade openness: fresh evidence from BRICS countries. Renew. Sustain. Energy Rev. 39, 14–23.

[13] Muhammad, B., 2019. Energy consumption, CO_2 emissions and economic growth in developed, emerging and Middle East and North Africa countries. Energy 179, 232–245.

[14]   McGowan, LD, Peng, RD, Hicks, SC. Design Principles for Data Analysis. Journal of Computational and Graphical Statistics 32. 754-761

[15]   Liu, Y., Yang, Y., Pan, F., Xue, D., 2022. A conformable fractional unbiased grey model with a flexible structure and its application in hydroelectricity consumption prediction. J. Clean. Prod. 367, 133029

[16]   Gradín C, Wu B. 2020.Income and consumption inequality in China: a comparative approach with India. China Econ Rev; 62: 101463.

[17]   Wang M, Feng C. 2019. Technological gap, scale economy, and China's industrial energy demand. J Clean Prod; 236:117618.

[18]   Murshed, M., Saboori, B., Madaleno, M., Wang, H., Doˇgan, B., 2022. Exploring the nexuses between nuclear energy, renewable energy, and carbon dioxide emissions: the role of economic complexity in the G7 countries. Renew. Energy 190, 664–674.