

# Research on Carbon Emissions of Industrial Clusters in China

Anning Ye, Yuguo Ji\*, Min Zhang

School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu, China

\*Corresponding author: Yuguo Ji (Email: 2723946533@qq.com)

**Abstract:** At present, the global warming problem is becoming more and more serious, and effective carbon emission reduction is urgent, and the cooperation between industries within a specific supply chain can provide a new method to reduce emissions. Whith 2017 year as the research period, 30 industrial sectors in China as the research object, using the new method proposed by Kanemoto et al. to identify high carbon emission industrial clusters. Combined with modified normalized cut function, we find out high carbon emission industrial clusters among 30 industrial sectors from the supply chain perspective with multiple clustering methods, and based on this, the relative position of each industrial sector in the industrial chain is studied through minimum spanning tree to find the key industrial chain. The results show that the clustering effect performs best at  $k=7$ , where cluster 1 accounts for 89% of the total carbon emissions of all clusters, indicating that this industrial cluster has more potential for emission reduction compared with other industrial clusters and is the focus of future emission reduction efforts, while the upstream and downstream industrial chains with the construction industry as the core are the key industrial chains of this cluster as shown by the minimum spanning tree.

**Keywords:** Input-output analysis, Cluster analysis, Minimal spanning tree, Chains of industry.

## 1. Introduction

Along with the increased frequency of global climate disasters, which in turn threaten the survival and development of human beings, the issue of carbon emissions has been continuously emphasized in various countries, however, in the context of industrialization and urbanization process, there are interconnections between industrial sectors in trade exchanges and product exchanges, therefore, in the implementation of carbon emission reduction policies, if a single constraint on high carbon emission industrial sectors will not be conducive to the reduction of emissions in the whole industry system Schliephake et al. proposed that in the field of supply chain management, cooperation between industries is an important way to collectively reduce costs, energy consumption, and net environmental impact [2]. Wang proposed that the wider and denser the scope of cooperation among supply chain enterprises, the more beneficial to the overall supply chain emission reduction effect [3]. So if some high carbon emission industries, such as petroleum processing as well as electric power and heat, can be found from the perspective of upstream and downstream supply chain relationships by clustering methods to find high carbon emission industrial clusters, then if the key sectors in these important industrial chains are subject to industrial cooperation, they may be able to effectively break through short-term technical limitations and other deficiencies to achieve the effect of emission reduction.

Kagawa et al. conducted a pioneering study on clustering input-output analysis by applying graph theory to input-output analysis, identifying a series of important industrial clusters from Japanese input-output tables [4]. After proposing this graph partitioning method for input-output tables, Tokito et al. combined input-output clustering analysis and structural path analysis to well classify the global transportation equipment supply chain network into emission clusters [5]. Kagawa et al. extended to global analysis using

the WIOD database and identified important clusters of the global supply chain network associated with final demand countries [6]. However, their objective function does not consider the whole supply chain. Kanemoto et al. propose a new approach to identify high carbon emission industry clusters [1]. It takes into account the carbon emission linkage of upstream and downstream industrial sectors in the supply chain in clustering.

When domestic scholars study carbon emission industry clusters, although there is relevant literature pointing out that cooperation among industry chain members can achieve the effect of emission reduction, there is a lack of suitable methods to find the relevant industry clusters, and it has not been studied through the perspective of clustering. In this paper, we will identify high carbon emission industrial clusters through input-output tables of China, and on this basis, we will find the key carbon emission industrial chains in industrial clusters through minimum spanning tree to provide new objectives and ideas for emission reduction actions within the supply chain.

## 2. Research Methods

### 2.1. Carbon Emissions Model

Influenced by the industry linkages, this paper uses Leontief's proposed formula for complete carbon emissions as follows.

$$F = \hat{f}(1 - A)^{-1}y \quad (1)$$

Where,  $F$  is total carbon emissions vector, also known as complete carbon emissions.  $\hat{f} = (\hat{f}_i) = \begin{pmatrix} e_i \\ x_i \end{pmatrix}$ , is the diagonal matrix of direct carbon emissions intensity (direct carbon emissions per unit of total output), where  $e_i$  is the direct carbon emissions from energy consumption (calculated by the IPCC method),  $x_i$  is the total output of industry sector  $i$ ,  $A =$

$(a_{ij}) = (\frac{T_{ij}}{x_i})$  is the direct consumption factor matrix,  $y = (y_j)$  is the final demand column vector. In addition,  $(1 - A)^{-1} = L = (L_{ij})$  is the Leontief inverse matrix.

## 2.2. Carbon Emissions Matrix

The variation on the basis of Equation 1, the equation is as follows.

$$F1 = \hat{f}(1 - A)^{-1}\hat{y} \quad (2)$$

Where,  $\hat{f} = (\hat{f}_i) = (\frac{e_i}{x_i})$  is a direct carbon emission intensity diagonal matrix.  $\hat{y} = (\hat{y}_j)$  is the final demand diagonal matrix.

## 2.3. Methodology for Clustering of Carbon Emission Industry Clusters

### 2.3.1. "Modified" Normalized Cut Function with Input-Output Analysis

Kagawa et al. failed to find high carbon emission industry clusters by considering the upstream and downstream relationships of industry sectors when clustering [4]. Kanemoto reworked the minimization normalized cut function from a directed graph perspective and used it as an objective function as follows [1].

$$\min Ncut_{dir}(C_1, \dots, C_m) = \sum_{r=1}^m \frac{cut_{out}(C_r)}{\sum_{i \in C_r} d_i^{out}} + \frac{cut_{in}(C_r)}{\sum_{j \in C_r} d_j^{in}} = \sum_{r=1}^m M(C_r) \quad (3)$$

Where,  $M(C) = \frac{cut_{out}(C)}{\sum_{i \in C_r} d_i^{out}} + \frac{cut_{in}(C)}{\sum_{j \in C_r} d_j^{in}}$ , for  $d_j^{in} = \sum_{i=1}^n f_i L_{ij} y_j - f_i y_j$  and  $d_i^{out} = \sum_{j=1}^n f_i L_{ij} y_j - f_i y_j$ , denote the carbon emissions achieved through the reverse linkage of downstream industry  $j$ , and the positive linkage of upstream industry  $i$ .

Then,  $cut_{in}(C_r) = \sum_{i \in S} \sum_{j \in C_r} f_i L_{ij} y_j - \sum_{i \in C_r} f_i L_{ij}^{C_r} y_j$  is the difference between the specific carbon emissions of sector  $i \in S$  and  $j \in C_r$  as a whole and the specific carbon emissions of sectors  $i$  and  $j$  within cluster  $r$ ,  $(I - A^{C_r})^{-1} = L^{C_r} = (L_{ij}^{C_r})$ .

Similarly,  $cut_{out}(C_r) = \sum_{i \in C_r} \sum_{j \in S} f_i L_{ij} y_j - \sum_{i \in C_r} f_i L_{ij}^{C_r} y_j$ , can be interpreted in a similar way.

### 2.3.2. Number of Clusters and Number of Industries Within Clusters

Since the focus of this study is to reduce carbon emissions through industrial cooperation, when using the modified normalized cut function as the objective function, this paper adopts Kanemoto's idea [1], which defines the maximum number of industries within a cluster as  $k$  and the number of clusters as  $m = \lceil n/k \rceil$ , determined by the number of industries  $n$  in the input-output table and the maximum number of industries  $k$  within each cluster. At the same time, the  $k$  value is too large, which will also be detrimental to the visualization of clusters and the extraction of key industry chains, so this paper specifies the maximum number of industries  $k$  as 5, 6 and 7, and finds the most suitable number of industries for important industry clusters from them.

### 2.3.3. Various Clustering Algorithms

Here in this paper, three alternative methods are proposed

based on Kanemoto, namely, greedy, hierarchical and shift-based clustering algorithms with a modified normalized cut function as the objective function [1]. However, the center-of-mass distance used in traditional clustering algorithms, such as the method of Ward [7], cannot be used in input-output clustering, while Kanemoto solves this problem by applying the modified normalized cut function to the clustering algorithm, with the final objective of obtaining a series of industry clusters  $C = \{C_1, C_2, \dots, C_m\}$ . There are two main methods for hierarchical clustering: aggregation method and splitting method. Also there are two methods for the movement-based algorithm: the objective function is normalized modified cut function and the maximization of carbon emissions within the cluster, respectively.

### 2.3.4. Intra-cluster Industrial Network Simplification — Minimum Spanning Tree

Here, the basic idea of Kruskal's algorithm is borrowed and modified, because the network in this study is a directed band-weighted graph, both the input-output relationship between industries exists in a directed manner, specifically in the process of industry production, the need to consume carbon emissions from the inputs of other industries, so in the process of selecting the optimal edges, it is necessary to take into account the directed nature of the construction network, and take the edges according to the edges from small to large, and merge the edges. This process ensures that no circle is formed after merging the edges, so it is also known as the circle avoidance method. The algorithm process is as follows.

- (1) First, sort the edges by weight from smallest to largest.
- (2) Take the edge with the minimum weight, and if 2 nodes in the edge are different components, then add the minimum spanning tree.
- (3) And so on until all edges are traversed
- (4) End condition of the algorithm: remaining edges = number of nodes - 1.

## 3. Data Source and Processing

### 3.1. Data Description

In this paper, the input-output table is used as the data base for sectoral division, and the input-output data involved in Equation 1 are obtained from the "2017 Input-Output Table" published on the official website of the National Bureau of Statistics, and the direct carbon emissions from energy consumption in Equation 1 are calculated by the IPCC method. Among them, "coal", "coke", "crude oil", "gasoline", "gasoline kerosene", "diesel", "fuel oil" and "natural gas". The carbon emission factors for each type of energy are determined based on data from the IPCC Greenhouse Gas Emissions Inventory.

### 3.2. Industry Sector Classification

Since the sectoral classifications in the input-output tables of each year and the sectoral classifications in the China Energy Statistical Yearbook of the corresponding year are not consistent, China's industrial sectors are combined into 30 industrial sectors considering the data matching problem. The results are shown in Table 1.

**Table 1.** Industrial sector classification

Code Name	Industrial Sector	Code Name	Industrial Sector	Code Name	Industrial Sector
S1	Agriculture, Forestry and Fisheries	S11	Petroleum processing and coking industry	S21	Instrument manufacturing
S2	Coal mining and washing industry	S12	Chemical industry	S22	Scrap waste and other manufacturing products industry
S3	Oil and gas extraction industry	S13	Non-metallic mineral products industry	S23	Metal products, machinery and equipment repair services
S4	Metal mining industry	S14	Metal smelting and rolling processing industry	S24	Electricity, heat production and supply industry
S5	Non-metallic ore and other mineral extraction industry	S15	Metal products industry	S25	Gas production and supply industry
S6	Food and tobacco industry	S16	General Equipment Manufacturing	S26	Water production and supply industry
S7	Textile industry	S17	Specialized equipment manufacturing	S27	Construction
S8	Textile, clothing, shoes, hats, leather and down and its system industry	S18	Transportation equipment manufacturing	S28	Transportation, storage and postal industry
S9	Wood processing and furniture manufacturing	S19	Electrical machinery and equipment manufacturing	S29	Wholesale, retail, accommodation and catering
S10	Paper printing and stationery manufacturing	S20	Communication equipment and other equipment manufacturing	S30	Other services

## 4. Carbon Emission Study of Industrial Clusters

### 4.1. Comparison of Multiple Clustering Results

#### 4.1.1. Evaluation Metrics

Since the focus of this study is to find out the maximum carbon emission in the industry cluster, this paper uses the sum of carbon emission within the cluster  $\sum_{r=1}^m \sum_{i \in C_r} f_i L_{ij}^{Cr} y_j$  as the evaluation index to compare the advantages and disadvantages of different algorithms to identify the high emission industry clusters, and if the sum of carbon emission in the cluster is higher, it indicates the better performance of the division of the supply chain network.

#### 4.1.2. Evaluation Results

The results of emissions within clusters by using different clustering algorithms targeting 30 industries in China are shown in Figure 1. The greedy algorithm performs the best in carbon emissions within  $k=5$ ,  $k=6$  clusters with the total sum of cluster carbon emissions of 6.8 billion tons as well as 7.6 billion tons of CO<sub>2</sub> respectively, however, the long computation time makes it difficult to calculate from 30 or more sectors. The second algorithm, based on the movement, is faster in terms of computational efficiency and performs best in clustering at  $k=7$ , with a total sum of cluster carbon emissions of 8.3 billion tons of CO<sub>2</sub>, and the second best performance at  $k=5$  and 6. When increasing the maximum number of industries  $k$  from 5 to 6, it can be considered, using the greedy algorithm, and when  $k$  is 7, the second one can be considered using the movement-based algorithm.

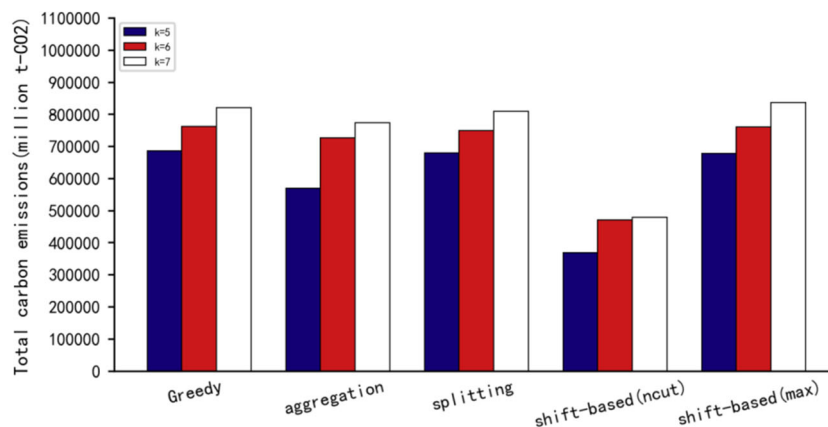


Figure 1. Clustering algorithm results show

### 4.2. Industry Chain Analysis in Industrial Clusters

From the results in the previous section, it is clear that the clustering effect is better as the sum of carbon emissions within a cluster increases with increasing  $k$  values when  $k$  is specified as 5, 6, and 7. Next, the modified minimum

spanning tree algorithm is used to extract the main network structure in the industrial network constructed by the inverse matrix of carbon emission values (Equation 2), and the results are shown in Figure 2. The main industrial chains extracted for each industrial cluster, as well as the connected edges with the smallest weights and all network nodes in the network are retained through the minimum spanning tree, and the

connected edges with large weights are dropped, where the network nodes represent each cluster internal industrial sectors, and the connecting edges of the nodes represent the inverse of the carbon emission values generated between two industrial sectors through supply chain transactions.

According to the number of connections between each node and other industrial sectors, it can be found that the more connections with other nodes, the more important the industrial sector is in the industrial group, so the core industrial sector of each industrial cluster can be determined. "textile, clothing, footwear, leather and down" (8), "transportation, storage and postal services" (28), "metal and mineral products" (13), etc. These industry sectors have the largest number of associated nodes in each industry cluster and are located at the center of the industry clusters and at the very center of the industry backbone network, and are also the hub nodes of each industry cluster. However, the core node in

an industry cluster does not necessarily mean that the sector has a large carbon emission value, because the clustering approach in this paper takes into account the upstream and downstream situation in the supply chain, and it is more likely to consider the implied carbon emission value brought by the demand from the upstream and downstream supply chain, and the larger carbon emission generated by the cooperation between industries is the focus of the study, for example, the construction industry is the core sector in cluster 1. For example, the construction industry is the core industry sector in Cluster 1, which does not have high carbon emissions in its own production, but the use of more intermediate products drives the carbon emissions of other industries in the supply chain to be higher, and measuring carbon emissions from a production perspective obviously ignores the "implicitly high carbon" characteristic of the construction industry.

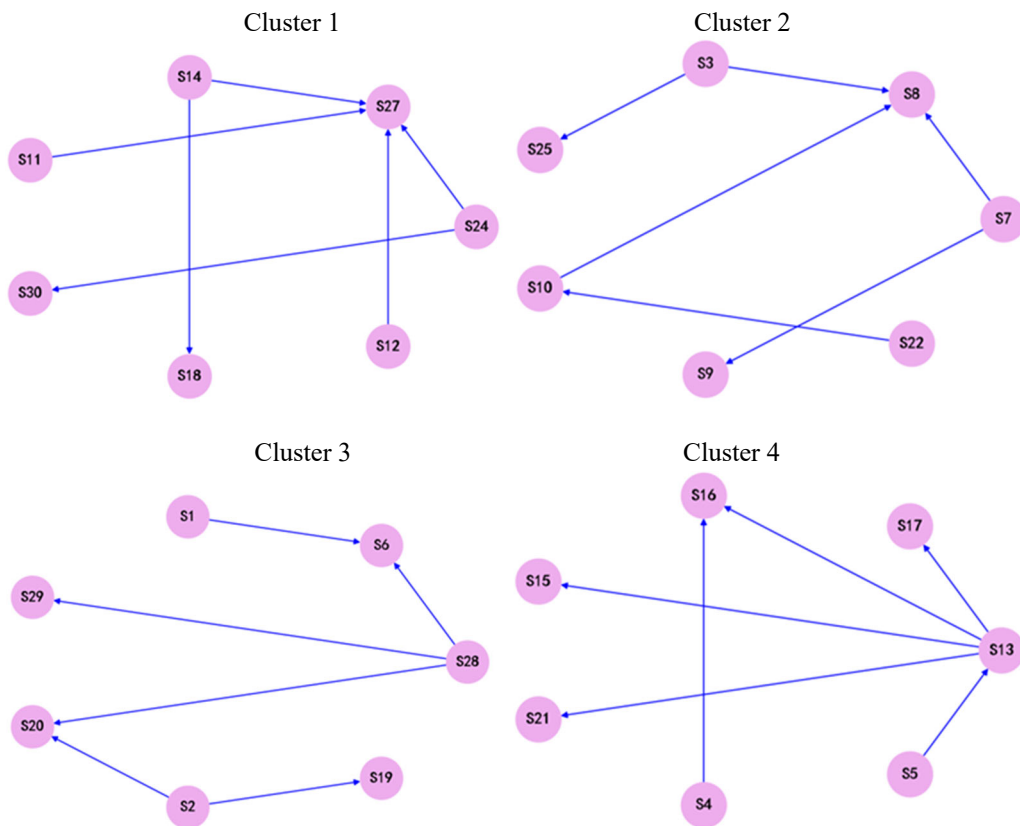


Figure 2. Minimum spanning tree for each industry cluster

Among them, cluster 1 accounts for 89% of the total carbon emissions of all clusters, indicating that this industry cluster has more potential for emission reduction compared to other industry clusters and is the focus of future emission reduction efforts. Because of the length of the article and other reasons, this paper only visualizes the minimum spanning tree results of cluster 1 for the next step of analysis, and studies the key industrial chains that exist in the important industrial clusters. The Sankey diagram can better observe the carbon emission flow situation within the industrial clusters and the upstream and downstream relationships of each sector in the industrial chains, and the results are shown in Figure 3. The industrial sectors of electricity, heat production, and oil processing are on the left, and these industries are located in the upstream

position of the supply chain in this industry cluster, and are the main source of carbon emissions for the downstream industries in this industry chain, while on the right are other service industries, construction, transportation equipment and other sectors, which are located in the downstream position of the supply chain in this industry cluster, and are also the final destination of carbon emissions in this supply chain. The thickness of the lines also indicates the amount of carbon emissions generated from one industry sector to another, representing the carbon emissions from and to different industry sectors. Therefore, the industry chain with the construction industry as the core should be the key industry chain to focus on when implementing the emission reduction policy.

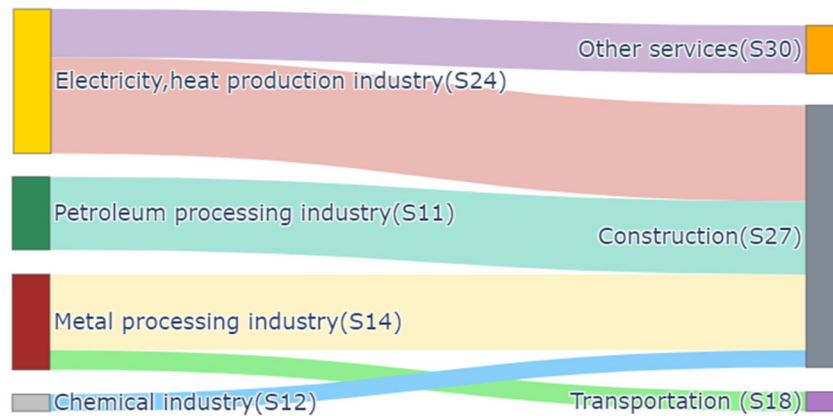


Figure 3. Cluster 1 Core Chain Carbon Emission Flows

## 5. Conclusions and Recommendations

### 5.1. Conclusion

In this paper, 30 industrial sectors in China are studied and a new method for identifying high carbon emission industrial clusters proposed by Kanemoto is used. Combining the modified normalized cut function to find high carbon emission industrial clusters from 30 industrial sectors and determine the optimal industrial clusters by relevant evaluation indexes, and on this basis, the relative position of each industrial sector in the industrial chain is studied by minimum spanning tree to find the key industrial chain, and the conclusions are as follows.

The results of multiple clustering algorithms based on input-output are seen: under the case of specifying the number of industries in each industry cluster as 5, 6 and 7 respectively, it is found that as  $k$  increases, the higher the sum of carbon emissions within the cluster, the better the industry cluster clustering effect. And the second algorithm based on the movement, the clustering effect performs best at  $k=7$ . Meanwhile, the results of the minimum spanning tree reveal that the core points of each industrial cluster are the industrial sectors of construction, textile, transportation, and metal and mineral making. At the same time, since the carbon emission of cluster 1 accounts for 89% of the total carbon emission of the cluster, it indicates that this industry cluster has more potential for emission reduction compared with other industry clusters, so it is necessary to further study the key industry chains for this industry cluster, and the construction industry sector is located in the core of this industry cluster, and the results of the minimum spanning tree of the Sankey diagram show that the carbon emission mainly comes from the electricity, heat production and supply industry, petroleum processing, etc. The upstream industry sectors bring higher carbon emission consumption.

### 5.2. Recommendation

The above study shows that carbon emissions should not only focus on the carbon emissions generated by the industrial sectors themselves, but also take into account the diffusion effects between sectors. And the extraction of key industrial chains in high carbon emission industrial clusters can more clearly see the linkage effect of the whole carbon emission system, the material flow between industrial sectors, and also this paper argues that these industries have the potential for effective emission reduction cooperation, and suggests that action should be targeted there to coordinate the development

of practical and reasonable synergistic emission reduction policies. Based on the above research and conclusions, this paper makes the following recommendations.

(1) Based on the industrial sector carbon emissions, for energy-consuming industrial sectors, such as the power and heat industries, they are the largest production-based emission sectors. For such industries, to implement effective regulations on the production side of the industry and improve the energy structure, it is necessary to integrate the sectoral resources of the industry, give full play to its scale advantages, replace coal with clean energy as the main fuel for power generation, and at the same time For the construction and service industries, since the production process uses a large number of high-energy products, the carbon emissions of these industries should be monitored in the intermediate stages.

(2) Within high carbon emission industrial clusters, the various industrial sectors that are stakeholders in the supply chain should consider how to reduce carbon dioxide emissions associated with their product systems through technology and product improvement and supply chain management, and provide a framework for industrial cooperation to reduce emissions, as doing so will maximize their carbon reduction potential

(3) Provide data to support the development and implementation of collaborative emission reduction policies through key industry chains identified by minimum generation trees. Emission reduction efforts should focus on key industrial chains in high carbon emission industrial clusters, pay attention to the dual monitoring of the source and intermediate sectors of the industrial chain, conduct comprehensive management of each industrial sector upstream and downstream, pay attention to the functions of each sector in the industrial chain and key links in the industrial chain, and propose energy-saving and emission reduction strategies in a targeted manner.

## Acknowledgment

This work is supported by scientific research project of Anhui university humanities and social sciences (Grant No: SK2021A0228).

## References

- [1] Kanemoto, K., Hanaka, T., Kagawa, S., Nansai, K. Industrial clusters with substantial carbon-reduction potential [J]. *Economic Systems Research*, 2019, 31(2):248–266.

- [2] Schliephake, K., Stevens, G., Clay, S. Making resources work more efficiently – the importance of supply chain partnerships [J]. *Journal of Cleaner Production*, 2009, 17(14):1257–1263.
- [3] Wang, Y., Long, C. Dual-domain emission reduction cooperation in three-tier supply chain under carbon trading policy[J].*Journal of Systems Management*, 2019,28(04):763-770.
- [4] Kagawa, S., Suh, S., Kondo, Y., Nansai, K. Identifying environmentally important supply chain clusters in the automobile industry [J]. *Economic Systems Research*, 2013,25(3):265–286.
- [5] Tokito,S., Environmentally-Targeted Sectors and Linkages in the Global Supply-Chain Complexity of Transport Equipment [J]. *Ecological Economics*, 2018,150:177–183.
- [6] Kagawa, S., Suh, S., Hubacek, K., Wiedmann, T., Nansai, K., Minx, J. CO2emission clusters within global supply chain networks: Implications for climate change mitigation [J]. *Global Environmental Change*, 2015,35:486–496.
- [7] Ward, J. H. Hierarchical Grouping to Optimize an Objective Function [J]. *Journal of the American Statistical Association*, 1963, 58(301):236–244.