Advances in Social Science, Education and Humanities Research (ASSEHR), volume 75
2016 International Seminar on Education, Innovation and Economic Management (SEIEM 2016)

Cluster Analysis of Energy Diversification Patterns in Different Regions of China

Li Xiaodi,Sun Shipeng,Yang Xincai
Harbin Engineering University
Harbin, China
958578783@qq.com

Abstract—**Due to the diversity of fundamental situation among provinces in China, there are visibly regional differences in economic development level, resource endowments, and technological development level. For this reason, the energy diversification shows a big diversity on development model. This paper, based on 1991-2013 energy consumption structure data of 30 provinces and cities, applying cluster analysis method of panel data, classify the changing energy structure pattern among target provinces and cities, finally summarizes ten typical types of energy diversification patterns.**

Keywords—energy structure; energy diversification pattern; cluster analysis

I. INTRODUCTION

China is a country covering vast territory; districts with different geography features, social economic and cultural environment shall form their unique pattern of energy diversification. In order to make a deep understanding of energy diversification patterns in different regions, we divide the energy consumption structure acquired from targets into groups by the method of cluster analysis.

At present, most cluster analysis is based on two-dimensional cross-section data, relatively analysis referring 3D panel has less applications. Therefore, some scholars have developed the method and application of cluster analysis for 3D panel data. Yinguo Li, Xiaoqun He(2010) [1]took the panel data feature of panel and time series into consideration, constructed the distance function and Ward clustering algorithm for measuring similarity. Finally, they achieved clustering analysis on panel data, and had a good test from the application. Lifeng Wu, Sifeng Liu (2013) [2] designed a panel data clustering analysis method to reflect the association degree of panel data by using the 3D grey convex relation degree, and found that this method was able to achieve a good clustering effect on the research objectives through actual test. Qunwei Wang, Shuangying Wang, Ze Cao (2014) [3] applied the self-organizing competitive network algorithm to the cluster analysis of panel data, and obtained a good application effect through the test of actual cases.

On the basis of three methods of of cluster analysis of panel data mentioned above, after considering the need of research content, we select the panel data clustering method designed by Sifeng Liu and Lifeng Wu. The breakthrough point is that this paper, which adopts clustering analysis for panel data, makes a deeper research on the basis of the clustering analysis of cross-sectional data. The results obtained from the information contained in the panel data of urban energy consumption structure gets closer to the actual situation and provides an accurate and effective method and technical support for a better analysis of typical models and characteristics of energy diversification.

II. METHOD OF PANEL DATA CLUSTERING FOR THREE-DIMENSIONAL GRAY CONVEXITY RELATION

The method of panel data clustering based on 3D gray convex relation is as follows:

The first step is to transform the panel data waiting for clustering into the behavior matrix of the object.

The second step is to average the original data so as to remove the dimension effect of different indexes, so

$$x_i'(s,t)=\frac{x_i(s,t)}{\overline{x}(s,t)'},\quad \overline{x}(s,t)=\frac{1}{L}\sum_{1}^{L}x_i(s,t)$$

The behavior matrix of the object i is changed into

$$x_i'=\begin{bmatrix} x_i'(1,1) & x_i'(1,2) & \cdots & x_i'(1,n) \\ x_i'(2,1) & x_i'(2,2) & \cdots & x_i'(2,n) \\ \vdots & \ddots & \vdots & \vdots \\ x_i'(m,1) & x_i'(m,2) & \cdots & x_i'(m,n) \end{bmatrix}$$,Where i = 1,2, ⋯,L.

The third step is to calculate the convexity of each point in the panel data, and then calculate the degree of correlation between any two objects, and compute an upper triangular correlation degree from the calculated association data, as shown in the following matrix:

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1L} \\ & \gamma_{22} & \cdots & \gamma_{2L} \\ & & \ddots & \vdots \\ & & & \gamma_{LL} \end{bmatrix}$$

The fourth step is to set a reasonable threshold r of correlation, according to the correlation matrix calculated in the third step, and get the cluster analysis results of panel data. In practical applications, we take the appropriate critical value of correlation is generally 0.5 to 1, that is $r \in [0.5,1]$.

When $\gamma_{ij} > r$, we think i and j are in the same cluster group, according to this method, traversing the correlation matrix, we get the final panel data clustering analysis results.

III. PANEL DATA CLUSTER ANALYSIS RESULTS

Because there isn't any cluster analysis application for panel data, we choose MATLAB and EXCEL to collate data and compute the data matrix, finally achieve empirical process and get results. According to the four steps mentioned above, we get the correlation matrix of clustering analysis for 30 provinces as shown in TABLE I. By setting different correlation critical value r, we assign values to r as:0.7,0.8,0.83,0.84,0.85,0.88, then get the result which is greater than the critical value, finally get the clustering results as shown in Table II.

TABLE I. TYPICAL ENERGY CONSUMPTION STRUCTURE OF PANEL DATA CLUSTER ANALYSIS RESULTS

Bei Jing	Tian Jin	Hei Bei	Shan Xi	Neimenggu	Liao Ning	Ji Lin	Heilongjian	Shang Hai	Jiang Su	Zhe Jiang	An Hui	Fu Jian	Jiang Xi	Shan Dong	He Nan	Hu Bei	Hu Nan	GuangDon	GuanGXi	Hai Nan	ChongQing	SiChuan	Gui Zhou	YunNan	Shan Xi	Gan Su	Qing Hai	Ning Xia	Xin Jiang	
1	0.71	0.67	0.66	0.66	0.71	0.69	0.67	0.72	0.67	0.67	0.66	0.65	0.66	0.68	0.65	0.67	0.65	0.7	0.64	0.49	0.66	0.64	0.66	0.65	0.7	0.72	0.68	0.66	0.65	Bei jing
	1	0.73	0.71	0.71	0.81	0.76	0.78	0.8	0.74	0.74	0.71	0.7	0.71	0.75	0.73	0.72	0.69	0.79	0.69	0.45	0.69	0.66	0.72	0.73	0.78	0.81	0.66	0.73	0.65	Tian jin
		1	0.85	0.84	0.73	0.82	0.71	0.75	0.82	0.82	0.88	0.78	0.87	0.81	0.85	0.85	0.85	0.75	0.81	0.47	0.66	0.62	0.86	0.85	0.74	0.77	0.62	0.8	0.62	Hebei
			1	0.9	0.72	0.79	0.7	0.75	0.81	0.82	0.87	0.76	0.83	0.78	0.79	0.83	0.81	0.76	0.8	0.47	0.67	0.62	0.88	0.84	0.74	0.76	0.61	0.77	0.61	Shanxi
				1	0.73	0.79	0.7	0.75	0.82	0.8	0.87	0.76	0.83	0.78	0.8	0.83	0.8	0.76	0.81	0.47	0.65	0.62	0.84	0.82	0.74	0.77	0.61	0.8	0.61	Neimenggu
					1	0.77	0.79	0.77	0.76	0.75	0.73	0.71	0.72	0.76	0.75	0.73	0.71	0.77	0.69	0.45	0.7	0.68	0.74	0.74	0.76	0.8	0.66	0.73	0.67	Liao ning
						1	0.74	0.78	0.87	0.83	0.79	0.78	0.8	0.83	0.83	0.83	0.79	0.77	0.75	0.46	0.68	0.63	0.79	0.79	0.76	0.81	0.63	0.81	0.62	Jilin
							1	0.74	0.75	0.74	0.7	0.7	0.7	0.74	0.71	0.71	0.69	0.74	0.67	0.45	0.73	0.67	0.71	0.7	0.76	0.77	0.63	0.71	0.66	Heilongjiang
								1	0.78	0.78	0.76	0.73	0.76	0.76	0.74	0.76	0.73	0.82	0.71	0.47	0.68	0.66	0.73	0.73	0.8	0.83	0.67	0.74	0.66	Shanghai
									1	0.79	0.77	0.73	0.77	0.77	0.75	0.78	0.74	0.82	0.71	0.47	0.7	0.67	0.74	0.74	0.8	0.84	0.68	0.74	0.66	Jiang su
										1	0.83	0.83	0.84	0.83	0.78	0.84	0.82	0.83	0.78	0.47	0.71	0.65	0.78	0.78	0.78	0.79	0.63	0.8	0.63	Zhe jiang
											1	0.78	0.88	0.8	0.81	0.87	0.87	0.77	0.83	0.46	0.65	0.62	0.84	0.82	0.75	0.76	0.61	0.8	0.61	Anhui
												1	0.8	0.79	0.76	0.78	0.79	0.78	0.8	0.45	0.66	0.61	0.74	0.74	0.73	0.74	0.61	0.77	0.61	Fujian
													1	0.8	0.83	0.88	0.89	0.78	0.81	0.47	0.66	0.62	0.81	0.81	0.75	0.76	0.61	0.83	0.61	Jiang xi
														1	0.82	0.83	0.78	0.77	0.73	0.46	0.68	0.63	0.79	0.78	0.74	0.78	0.62	0.77	0.63	Shan dong
															1	0.83	0.81	0.73	0.76	0.46	0.66	0.62	0.82	0.83	0.73	0.76	0.61	0.79	0.61	Henan
																1	0.75	0.7	0.7	0.6	0.64	0.63	0.7	0.71	0.69	0.7	0.63	0.73	0.63	Hubei
																	1	0.76	0.8	0.46	0.64	0.6	0.79	0.79	0.74	0.74	0.61	0.81	0.61	Hunan
																		1	0.74	0.46	0.69	0.66	0.73	0.72	0.78	0.81	0.65	0.75	0.64	Guangdong
																			1	0.46	0.64	0.6	0.8	0.79	0.72	0.72	0.6	0.76	0.59	Guangxi
																				1	0.47	0.47	0.46	0.46	0.48	0.46	0.45	0.46	0.47	Hai nan
																					1	0.75	0.67	0.66	0.71	0.7	0.64	0.65	0.71	Chongqing
																						1	0.63	0.63	0.66	0.68	0.63	0.61	0.69	Si chuan
																							1	0.91	0.72	0.74	0.61	0.75	0.62	Gui zhou
																								1	0.72	0.74	0.62	0.75	0.62	Yun nan
																									1	0.79	0.68	0.73	0.65	Shanxi
																										1	0.66	0.76	0.64	Gansu
																											1	0.59	0.62	Qing hai
																												1	0.6	Ning xia
																													1	Xin jiang

TABLE II. PANEL DATA CLUSTER ANALYSIS GROUP RESULTS

Groups	Group members (province, city, district)
Group One	Inner Mongolia, Shanxi, Guizhou and Yunnan
Group Two	Hubei, Hunan, Anhui, Jiangxi, Hebei, Henan, Ningxia
Group Three	Jilin, Shandong, Shaanxi, Jiangsu, Zhejiang
Group Four	Heilongjiang, Liaoning, Tianjin, Gansu, Shanghai, Guangdong
Group Five	Fujian, Guangxi
Group Six	Sichuan, Chongqing
Group Seven	Beijing
Group Eight	Qinghai
Group Nine	Xinjiang
Group Ten	Hainan

The first group of energy diversification patterns is clustered in the southwest and the north; the second group of energy diversification patterns is clustered in the central region; the third group of energy diversification patterns is clustered in one province in northeast China, three provinces in East China and one province in Northwest China. The fourth group of energy diversification patterns is located in two provinces in northeast, one in northwest and one in southeast. The fifth group of energy diversification patterns is distributed in South China and the sixth group of energy diversification patterns is clustered in the southwestern provinces. The seventh group of energy diversification patterns is only in Beijing. The eighth type of energy diversification patterns is Qinghai model. The ninth type of energy diversification patterns is Xinjiang model. The tenth energy diversification patterns are Hainan model. It can be seen from the spatial distribution that groups have a strong correlation in geographical locations, a phenomenon that should be related to that similar energy resources exist in the relevant area, so members of one group have a similar resource endowment and show a similar pattern of energy diversification. Of course, the factors that affect the diversification of energy resources are far more than resource endowment; other factors such as technology, environment, society, population also make significant impact.

IV. Comparative Analysis of Typical Energy Diversification Patterns

According to the cluster analysis results of the above table, the thirty sample provinces and cities in this paper are divided into ten different groups. Based on the analysis of the characteristics and changes of each group data, ten typical energy diversification mode and characteristics are summarized.

Firstly, we make definition of six small concepts. Proportion between 80% to 100% is called a Absolutely Dominant(AD), between 65% and 80% called it Strong Dominant(SD), 50% to 65% called Weakly Dominated(WD), 30% to 50% is called Absolute Added(AA), 15% to 30% as Strong Complement(SC), added 0% to 15% as Weakly Complement(WC).

TABLE III. CHARACTERISTICS OF TEN TYPICAL MODELS OF ENERGY DIVERSIFICATION

Category	Consumption ratio	Group characteristics
The first category	Coal: 85% -95%; Oil, Natural gas: close to 0%; Electricity: close to but less than 10%	Coal AD
The second category	Coal:70% -80%; Electricity: close to but higher than 10%; Crude oil: close to but less than 10%; Natural gas :gradually increased from 0% to 2% -4%	Coal SD Crude oil WC Electricity WC
The third category	Coal :60% -70%; Crude oil: close to 20%; Electricity: close to 10%; Gas:0% gradually increased to 2% -5%	Coal SD Crude oil SC Electricity WC
The fourth category	Coal :50% -60%; Crude oil: about 30%; Electricity: increased to 10% -15%; Natural gas :from 0% -3% to 4% -9%	Coal WD Crude oil SC Electricity WC
The fifth category	Coal :from 85% -90% absolute lead into 55% -60% weak leading; Crude oil :increased from single digits to 15% -20%; Electricity :increased from just over 10% to 20%	Coal AD to WD Crude oil WC to SC Electricity WC to SC
The sixth category	Coal :from 80% -85% dominant to 60% -70% dominated; Crude oil: remained close to 0%; Natural gas and Electricity from slightly lower than nearly 10% to 15% -20% and tend to 20%	Coal AD to WD Electricity WC to SC
The seventh category	Coal:69% ;Crude oil: 33% Electricity:7% ;Natural gas :1% Gradually developed into a very balanced energy structure, the consumption ratio of coal, crude oil, natural gas, electricity is about 20%	Coal AD to WD Crude oil AC to SC Electricity WC to SC
The eighth category	Coal : 77%, Crude oil :10%, Electricity:12%, Natural gas :about 1%, Gradually transformed into Coal:absolutely 40%, Crude oil: 10% Natural gas :about 20%, Electricity: about 30%	Coal SD to AC Crude oil WC Electricity WC to SC
The ninth category	Coal : 61%, Crude oil : 32%, Electricity and Natural gas: 7% Gradually evolved into Coal : 50%, Crude oil : 20%, Electricity and Natural gas: about 10%.	Coal WD Crude oil SC Electricity WC
The tenth category	Coal: 74% ; Electricity: 26% ; Crude oil and Natural gas :0% Gradually evolved into Crude oil :about 40% Coal and Natural gas:about 25% Electricity: about 10%	Coal SD to SC Crude oil WC to AC Electricity SC to WC

V. CONCLUSION

In this paper, we cluster the energy diversification patterns of 30 provinces and cities by cluster analysis of time series data and panel data, and compare and analyze the characteristics of diversification patterns of different groups, and get some typical energy diversification models, which has a certain reference value for the development of various regions in our country and the energy diversification mode which adapts to its own conditions.

ACKNOWLEDGMENT

I would like to express my gratitude to all those who helped me during the writing of this paper. With the help of my supervisor Professor Li, I have overcome difficulties. I do appreciate her patience, encouragement, and professional instruction. What's more, I want to thank my thesis advisor, who kindly gives me a hand during my paper writing. At last, comments and suggestions are most welcome.

REFERENCES

[1] Cingguo LI, Xiaoqun HE.Cluster Data Clustering Method and Its Application [J]. Statistical Research, 2010, 27(9): 75-81. In Chinese

[2] Lifeng WU, Sifeng LIU.Cluster Data Clustering Method Based on Gray Convexity Degree and Its Application [J] .Control and Decision, 2013, 28 (7): 1032-1036, 1045. In Chinese

[3] Shuangying WANG, Qunwei WANG, Ze CAO.Multi - index Panel Data Clustering Method and Its Application - Taking Industry Primary Energy Consumption Panel Data as an Example [J].Mathematical Statistics and Management, 2014, 33 (1): 42-49. In Chinese

[4] Xianguo Li. Diversification and localization of energy systems for sustainable development and energy security[J]. Energy Policy, 2005, (33): 2237-2243. In Chinese

[5] Lior N. Sustainable Energy Development: the Present (2011) Situation and Possible Paths to the Future[J]. Energy, 2012, 43(1): 174-191.

[6] Stirling Andy. A general framework for analysing diversity in science, technology and society[J]. Journal of the Royal Society Interface, 2007, 4(15): 707-719.