

The Cluster Analysis about Tourism Benefit in Hainan

He Lifang

Depart. Of Mathematics and Statistics
Hainan Normal University
Haikou, Hainan
634224845@qq.com

Zhang Chengyi

Depart. Of Mathematics and Statistics
Hainan Normal University
Haikou, Hainan
chengyizh@sina.com

Chen Yijuan

Depart. Of Mathematics and Statistics
Hainan Normal University
Haikou, Hainan
867146636@qq.com

Abstract—Based on data from the hainan statistical yearbook between 2002 and 2012, this paper makes fuzzy cluster analysis of eighteen cities of Hainan by introducing FCM algorithm, and divides these cities into four categories. We compare and analyse the data of cities. The main conclusions are as follows: 1)The tourism benefit of the city is not proportional to the development of the city; 2)Wangning is the preferred place to settle for the tourists relative to other cities such as Qionghai. 3)The cost of hospitality is about twice as much as that of travelling. 4)The foreign tourists are inclined to choose Sanya and Haikou to settle. 5)The gap of tourist benefit between eastern line and western line is very big. At the end, countermeasures and suggestions are offered for these phenomena and their reasons.

Keywords—tourism benefit; FCM algorithm; fuzzy cluster analysis; analyse ; suggestions

I. INTRODUCTION

Hainan located in the southernmost of China is a tropical tourist island and is one of the most popular tourist resort . In particular,since the construction of Hainan international tourist island turned into a national strategy,the tourism industry of Hainan developed rapidly. It is significant to classify these cities, evaluate the development of these cities and put forward targeted policy and suggestion on the basis of past data since the tourism benefit of cities is one of factors of evaluating the development of tourism.

Cluster analysis is a mathematical method of classifying objects in according to their different characteristics, degree of intimacy and similarity. But in the real world, the bound of relationship among objects is not trenchant, we call it fuzzy relationship.It is more natural and realistic to classifying using fuzzy mathematical method. Fuzzy C mean value is a method of classification to find the best cluster using fuzzy ISODATA algorithm on condition that the number of cluster is given. According to Hainan statistical yearbook, among these important statistical indexes,there are five indexes can reflect the condition of tourism benefit. These

are:the number of demostic passengers served by tourist hotel, the number of overseas passengers served by tourist hotel, demostic tourist income, tourist income of foreign exchange, average occupancy of hotel. This paper classifys 18 cities in Hainan based on their 5 characteristic indexes from 2001 to 2011. Not only can we analyze present situatilon of tourism by classifying cities about tourism benefit, but also we can find their common chatacteristic or difference by reseaching the data in same class or different class,and we can find the factor by comparing the data of cities which are on different classes but a part of data are semblable . This paper analyzes the factor ,put forward relevant policy,measures and suggestions aimed at above-mentioned three phenomena.

II. FCM CLUSTERING ALGORITHM AND MATALAB REALISATION

FCM algorithm is a fuzzy clustering method based on objective function and was put forward and improved by Bezdek . FCM algorithm is one of the most extensive fuzzy clustering algorithm since it is simple and its operating speed is fast.

Since the computing amount of fuzzy clustering is considerable, it is impossible to complete computing manually or by computing counter, and we need software programming to implement.Matlab is a professional mathematical software developed by MATHWORK company,and it has capability of science operations,data analysis and programming.There is toolbox of fuzzy logic in it and it is convinient to implement fuzzy clustering analysis of data.

III. CLUSTERING CITIES IN HAINAN ON TOURIST BENEFIT BY USING FCM ALGORITHM

A. Indexes and data

According to the statistical yearbook of Hainan from 2002 to 2012,we select five indexes. They are the number of domestic passengers served by hotels of tourism,the number of overseas passengers served by hotels of tourism,domestic income of tourism,foreign income of

tourism and the average occupancy of hotels. These five indexes as above are positively related to income of tourism. Since the data in statistical yearbook is the data of last year, and the statistical yearbook in 2013 has been not published, in fact the data is the statistic of 2002 to 2011, but a part of data of 2013 has been published, we supplement the data of 2012 by fitting of time series. In addition, a part of data in statistical yearbook are gone, for example the data of occupancy of hotel in 2002 has gone, so we get the data by fitting the data of 2003 to 2008. Because constructing the international tourism island in 2009, the development of tourism was faster since then, it is reasonable that we select the data before 2008 to fit. A part of the data of foreign passengers has gone, in consideration of the data may be zero, we regard them as zero. Since the data of Baisha from 2002 to 2011 has gone, but there is the data of 2012. In consideration that the number of passengers may be seldom, we regard them as zero.

B. The pretreatment of original data

Because the dimension and the order of magnitudes of five indexes are different, the bigger order of magnitude can be highlight and the smaller order of magnitude can be eliminated in arithmetic process, so as to lack of uniform measure. In order to remove this influence, we should treat the data, and the corresponding transformation formula is:

$$u'_{ij} = \frac{u_{ij} - \min_{1 \leq i \leq n} \{u_{ij}\}}{\max_{1 \leq i \leq n} \{u_{ij}\} - \min_{1 \leq i \leq n} \{u_{ij}\}}, (j=1, 2, \dots, 5)$$

Obviously, $0 \leq u'_{ij} \leq 1$. We can get the original matrix of clustering by the pretreatment of original data.

C. Acquisition on the result of clustering

Fixed the number of clustering $c=4$, equivalently we cluster the 198 samples, and get the matrix of membership U . Lastly, we cluster the samples by using evaluating guidelines on the highest grade of membership. Perform the program of Matlab, and get the best result of clustering, see table below:

TABLE.I OUTPUT OF CLUSTERING CITIES OF HAINAN(NUMBER)

Categories	Serial number of cities contained in categories					
The first class	3	5	9	10	12	14
	15	23	24	27	28	
	29	30	32	33	41	
	45	46	47	48	50	51
	53	59	64	65	66	
	68	69	71	75		
	77	80	82	83	84	86
	87	89	93	94	95	
	98	100	101	102	104	
	105	106	107	111		
	112	113	116	118		
	119	120	122	123		
	124	125	129	130		
	131	134	136	137		
	138	141	142	143		

		147	148	149	152		
		153	154	155	156		
		158	159	160	161		
		165	166	167	170		
		171	172	173	174	176	
		177	178	179	183		
		184	185	188	189		
		190	191	192	195		
		196	197				
The second class	1	2	6	19	20	37	38
		42	55	56	60	73	78
		91	96	109	114	127	132
		145	150	163	168	181	186
The third class	4	7	8	11	13	16	17
		18	21	22	25	26	31
		34	35	36	39	40	43
		44	49	52	54	57	58
		61	62	63	67	70	72
		76	79	81	85	88	90
		97	99	103	108	115	117
		121	126	133	135	139	140
		144	151	157	162	169	175
		180	187	193	194	198	
The fourth class	74	92	110	128	146	164	182

Every city has 11 serial numbers, we appoint that the class which contain most serial numbers is the class which the city belong to. The final result is:

TABLE.II OUTPUT OF CLUSTERING CITIES OF HAINAN

Categories	Cities of the classes
The first class	Wuzhishan, Wenchang, Qionghai, Tunchang, Chengmai, Lingao, Danzhou, Dongfang, Qiongzong, Baoting, Linshui, Changjiang
The second class	Haikou, Wanning
The third class	Dingan, Ledong, Baisha
The fourth class	Sanya

According to the data, the data of Sanya's five indexes are bigger than other cities'. The data of Haikou and Wanning are bigger than other cities relatively. But the data of Baisha, Ledong and Dingan is smaller. We combine the situation of tourism and the data of statistical yearbook, and prove that the result is containing the true picture. The fourth class contains cities whose income of tourism is highest, the second class contains cities whose income of tourism is higher, the first class contains cities whose income of tourism is high, the third class contains cities whose income of tourism is low.

IV. ANALYSIS AND SUGGESTIONS FOR THE RESULT OF CLUSTERING

We can find some phenomena as follows by analyzing the result of clustering and the data of statistical yearbook.

First, the order on the basis of tourist development is not quite consistent with that of tourist benefit. We cluster

according to the tourist benefit ,not the development of tourism. To see the actual situation ,since Wenchang and Qionghai are near to Haikou,Baoting and Lingshui are near to Sanya.These four cities' level of development of tourism are higher than other cities,but Haikou and Wanning are clustered in same class,and these four cities are clustered in other class.In addition, we can find the five indexes of characteristic are related to the number of passengers or income of the hotels from tab.3.So the result of clustering can be closely connected with the income of tourism,but not closely connected with the tourist development of the cities.

TABLE.III STATISTICS OF 2010 (HAIKOU ,SANYA ,ETC.)

Cities	The number of demostic passengers served by tourist hotel	the number of overseas passengers served by tourist hotel
Haikou	4929500	132873
Sanya	7480900	400521
Wenchang	485300	5839
Qionghai	906000	13797
Wanning	2650800	36853
Baoting	135900	3927
Lingshui	819900	59
demostic tourist income	tourist income of foreign exchange	average occupancy of hotel
60.63	7161.514	60.6
92.02	21587.06	63.12
5.97	314.7071	58.21
11.14	743.623	58.25
32.61	1986.282	52.87
1.67	211.6552	50.73
10.08	3.179949	59.68

From:statistical yearbook of Hainan(2011),committee of tourism in Hainan

Second,the tourist line along Hainan's eastern line is usuallyHaikou → Qionghai → Wenchang → Wanning → Sanya.Wenchang and Qionghai which near to Haikou are the only way which travellers must pass by.Baoting and Lingshui are near to Sanya,and their number of passengers are more.These four cities should be the cities whose income of tourism is higher than others.But the result of clustering is not so.Investigating the cause,it may be because that Qionghai ,Wenchang,Baoting and Lingshui are the sight spots which passengers pass by,and most passengers don't spend the night at these spots but Wanning.But the five indexes are related to the number of passengers served by hotels,and it is also the occupancy of hotels. It is why Haikou and Wanning are clustered in the same class,but others are clustered in other class.

Third,we contrast Wanning and Qionghai. We contrast their data of first four indexes from Tab.3.In Fig .1,the data of Wanning are 2650800, 36853 , 32.60532, 1986.282, the data of Qionghai are 906000, 13797, 11.14397, 743.623.We can find that the data of Wanning is triple that of Qionghai.The passengers which

pass by Qionghai must pass by Wanning.Passengers just spend on spots' tickets.The rest are the expenses of meal and accommodation.So the expense of meal and accommodation are twice as much as the cost of tickets.

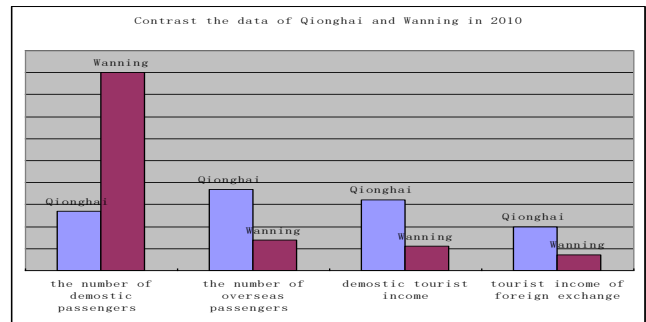


Figure.1 Contrast of data between Qionghai and Wanning

Forth,from the data of some indexes,we can find that the foreign passengers are inclined to descend on Sanya and Haikou.Because their tourist facilities are complete,normal and secure,and their service of medical and health are more sound.

Fifth,we can find that the gap between eastern income and western income is big from the data.For example,Qionghai and Wanning are located at eastern line.There are more people going to these places.But there are some cities located at western line,for example Danzhou,Lingao,Changjiang and Dongfang,there the passengers and income of tourism are little relatively.

TABLE.IV CONTRAST OF DATA BETWEEN THE EASTERN FRONT AND THE WESTERN FRONT

cities	the number of demostic passengers served by tourist hotel	the number of overseas passengers served by tourist hotel
Qionghai	906000	13797
Wanning	2650800	36853
Lingao	65100	0
Danzhou	501400	1320
Dongfang	116900	133
Changjiang	175100	0
demostic tourist income	tourist income of foreign exchange	average occupancy of hotel
11.14397	743.623	58.25
32.60532	1986.282	52.87
0.800742	0	81.08
6.167312	71.14462	58.59
1.437891	7.168359	67.02
2.153762	0	69.26

From:statistical yearbook of Hainan(2011),committee of tourism in Hainan

Based on the analysis on these situations above,we put forward some proposals as follows:

First,for Qionghai,Wenchang ,Baoting and Lingshui,we try to transform the passing by city to staying city,for example,we can add new scenic spots which can

constitute chain or groups of attractions with original scenic spots to lengthen the passengers' stranded time. Also we can lengthen the chain of industry, develop program of characteristic tour, for example, special shopping, constructing project of special tourism which is located at sports, anime and entertainment. The income of tourism can be raised by lengthening visiting time and letting passengers stay at the city. For example, in Qionghai, we can vigorously develop international convention and cultural creative industry, and build brand of international convention and cultural industry; In Qixianling, we can continue to perfect service facility and build a tourist resort features spa, forest tourism.

Second, develop scenic service project, for example, providing local special products, launching island tourism, offshore leisure sports, cruise and yacht. These can make more passengers be attracted and stimulated in spending. Perfect supporting services, achieve completed infrastructure and ensure supporting facility is completed to improve visitors strength of consumption.

Third, perfect the foreign passengers' service system in boutique scenery, such as providing travel service related to foreign language and culture, vigorously developing and cultivating international nursing travel product whose service content is medical care and senior citizen's service. These can appeal to foreign passengers and speed up the development of foreign markets and promote the development of Hainan's tourist industry.

Forth, narrow the gap between eastern tourism and western tourism. For example, develop special tourism at western spots, and develop unique tourist project different from eastern spots. Make the most of western advantage of tourist resource, such as harbor industry, high and new technology industry, resources of tropical rain forest, ecological resources and historical resources. Develop vigorously ecological tourism, industrial tourism and coastal tourism, develop explorative amusing project, and build seaside tourist living resort. For example, plan and construct Dongpo cultural park. Dig deep to and protect and promote the culture of Dongpo. Construct Bawangling to be a tourist resort whose integration with ecological tourism, leisure tourism and sound science education.

V. CONCLUDING REMARKS

In this paper, we introduce FCM algorithm to the clustering of tourist benefit, and get the optimal result of clustering by using Matlab to realize FCM algorithm. We find their characteristic by analyze the data of city in same class, find out their difference by contrasting and analyzing the data of cities in different class, and find the influencing factors by contrast the data of cities whose part of the component are semblable. We raised corresponding policy, measure and suggestions directed to these phenomena and cause. How to contact the tourist benefit and the level of tourist development and carry out corresponding evaluation and analysis are following work for us.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (Grant No. 71361008) and Hainan special social development technology (Grant No. 2011SFO3). Corresponding author: Chengyi Zhang.

REFERENCES

- [1] Zhang Min, Yu Jian. Fuzzy partitional clustering algorithms [J]. Journal of Software, 2004, 15(6):858-868(in Chinese).
- [2] Li Bonian. Weigh on cluster fuzzy C-Mean [J]. Fuzzy Systems and
- [3] Sun Xiaoxia, Liu Xiaoxia, Xie Qianru. The implementation of the fuzzy C-means clustering algorithm[J]. Computer Applications and Software, 2008, 25(3): 48-50(in Chinese).
- [4] Sun Jigui, Liu Jie, Zhao Lianyu. Clustering algorithms research[J]. Journal of Software, 2008, 19(1): 48-61(in Chinese).
- [5] Zhang Xueyan, Chen Chuanzhong, Hu Xiaohua. Forecast analysis of the tourist number in hainan by times series process methods[J]. Journal of Hainan Normal University(Natural Science), 2013, 26(1): 13-18(in Chinese).
- [6] Zhang Wenguo, Wang Guoli. Fuzzy partition model on groundwater flow zones[J]. Advances in Water Science, 2000, 11(2): 180-185(in Chinese).
- [7] Chen Shuili, Li Jinggong, Wang Xianggong. Fuzzy set theory and its application(the first Edition)[M]. Science Press, 2005. 94-118 (in Chinese).
- [8] Wang Fuyong, Dai Haohui, Pang Jianyu. Mathematical experiment course(Matlab Edition)[M]. Science Press, 2012. 101-116(in Chinese).
- [9] Lai Bing. A Study on the Market Competition Situation of Travel Service Routing Product Based on Cluster Analysis [J]. TOURISM TRIBUNE. 2006. 21(8):60-63(in Chinese).
- [10] Chen Ganghua. Tourism Penetration in Island-type Destination: A Case Study of Hainan and Its International Comparison[J]. TOURISM TRIBUNE. 2012. 27(11):72-80(in Chinese).