Zoning of Cultivated Land Intensive Use in Hubei Province Based on K-means Clustering

Cheng Qi

College of Arts and Communications, Wuhan Polytechnic University, Wuhan 430023, China

Abstract: Cultivated land intensive use is an important way of keeping agricultural sustainable development in rapid urbanization period. Traditional evaluative methods of cultivated land intensive use could not completely express the regional overall difference and partial spatial heterogeneity. This paper takes Hubei Province as study area, constructs the index system of cultivated land intensive use from three aspects of input intensity, use intensity and output efficiency, and uses K-means algorithm to zone the cultivated land intensive use level. The results show: (1) A good zoning result of cultivated land intensive use by using the characteristic of briefness and high-speed has been made. (2)There is a distinct spatial difference of cultivated land intensive use in Hubei Province. The result is zoned by western low-intensive zone, southern middle-intensive zone, eastern high-intensive zone and central fine-farming zone. The corresponding development strategy is also proposed according to the agricultural productive condition of each zone.

1. Introduction

Cultivated land is the most basic means of production in agricultural production, and is also a basic condition for ensuring national food security. Intensive use of cultivated land is not only a fundamental way to alleviate the contradiction between supply and demand of cultivated land, but also an important measure to promote social and economic development, [1]. Under the background of rapid industrialization and urbanization, the continuous expansion of the urban area has occupied the cultivated land, and the area of cultivated land in China has been declining continuously since 1980, [2]. Related research shows that in the past 1990-2000 years, farmland occupied by the whole country accounted for 45.96% of the total area of cultivated land in this period, and increased to 55.44%[3] in 2000-2010 years. At present, the research on cultivated land intensive degree is mainly concentrated on [4],. These studies have played a positive role in the intensive use of cultivated land, but the traditional method of measuring the spatial difference of intensive use of cultivated land can not fully reflect the overall regional differences and local spatial heterogeneity. The study of the intensive degree of cultivated land utilization is helpful to the scientific understanding of the spatial differentiation law of cultivated land use intensivism , and then the differential development, utilization and protection mode, in order to realize the division guidance of intensive use of cultivated land, to clear the emphasis of the protection of cultivated land, to optimize the pattern of agricultural development, to ensure the security of grain and to ensure the security of grain. Promoting sustainable development of social economy is of great significance.

Land zoning is one of the hotspots in land use research. At present, the research on land use zoning in China has been very rich, and the research area is diversified, and the research methods are becoming more mature.

Hubei province is vast in territory and rich in cultivated land resources. Jianghan Plain is an important food production base in China. On the other hand, Hubei is an important fulcrum in the strategic deployment of central China's rise. In recent years, the development of industrialization and urbanization has entered the white hot stage. Under the double pressure of cultivated land protection and rapid social and economic development, it is an effective way to improve the intension of intensive utilization of cultivated land. This paper takes the city and the state as the research unit, and uses the Kmeans algorithm to carry out the study on the intensive degree of cultivated land utilization in Hubei Province, which can provide the method support for the regional cultivated land use intensive degree management.

2. Research Areas and Data Sources

Hubei province is located in the middle of China, the middle reaches of the Yangtze River and the north of Dongting Lake. The geographical location is between 108 degree 21 '42 "~116 07' 50" E, 29 degree 01 '53 "~33 degree 06' 47" N, with the geographical advantage of the west, the South and the north and the four sides, the total area of the land is 185 thousand and 900 km2, among which the cultivated land area is 53 thousand and 200 km2. Most of the province is a subtropical monsoon humid climate with good natural conditions. The superior climatic conditions and rich cultivated land resources make Hubei an important agricultural production base in China. The comprehensive agricultural production capacity has been stable for a long time, and the output of main agricultural products is in the forefront of the country. Based on the social and economic statistics of Hubei province in 2013, this study takes 17 places, cities and states of Hubei Province as the research object to carry out the study on the area of cultivated land use intensive degree in Hubei province. The items involved include the agricultural employees in different cities and cities, the total power of agricultural machinery, the total input of agricultural chemical fertilizer, the total area of cultivated land, the sown area of crops, the effective irrigation area, the total grain output, the total agricultural output value, the total population of the region, the total land area and so on. The data used in this paper come from the 2014 Hubei statistical yearbook and the Hubei Rural Statistical Yearbook.

3. Research Methods

Clustering analysis is based on data similarity. The clustering analysis method can be used to analyze the similarity of multiple factors and their differences, and to study the differences between regions. It is an important method to quantitatively study the classification of geographical things and the problem of geographical areas. The common clustering methods include systematic clustering, dynamic clustering and fuzzy clustering. MacQueen put forward Kmeans algorithm in 1967. It is one of the effective algorithms widely used in many clustering algorithms of science and industry. The Kmeans algorithm is a non supervised real-time clustering algorithm. The data is divided into a predetermined number of K on the basis of minimizing the error function. The sample points in each cluster have high similarity, and the similarity degree between the clusters is low. The similarity calculation is based on the average of the sample points in a cluster. First, K objects are randomly selected. Each object initially represents the average or center of a cluster. For each remaining object, it is assigned to the nearest cluster according to its distance from each cluster center, and then the average value of each cluster is recalculated. This process is repeated until the criterion function converges. This algorithm tries to find the K partition that minimization of the criterion function value. When the result cluster is dense, and the difference between cluster and cluster is obvious, its effect is better. Processing large data sets, which is relatively scalable and efficient, because its complexity is O (nkdt), where n is the number of all objects, K is the number of clusters, D is the dimension of the data, and t is the number of iterations. Normally, k<<n, and t<<n. This algorithm often ends with a local optimal.

Construction of the index system of intensive degree of cultivated land use. Based on the principles of pertinence, direct, spatial heterogeneity, data availability and feasibility, this study, on the basis of the existing achievements, combined with the social and economic conditions of Hubei, the conditions of agricultural production and the utilization of cultivated land, from the three aspects of cultivated land input intensity, cultivated land use intensity and cultivated land output efficiency. The index system of intensive utilization of cultivated land is used to carry out the regionalization of cultivated land intensive utilization in Hubei province.

4. Results and Analysis

According to the statistical yearbook of Hubei province in 2014 and the statistical yearbook of Hubei rural statistics, we can get the 2013 Hubei provincial level and state agricultural employees,

the total power of agricultural machinery, the total input of agricultural chemical fertilizer, the total area of cultivated land, the sown area of crops, the effective irrigation area, the total grain and food output, the total agricultural output value and the regional general people. Based on the total area of mouth and land, the regional index of cultivated land intensive use in different cities of Hubei province is calculated. As there is a huge difference between the dimension of index score, it will have a negative effect on the area of cultivated land use intensity in Hubei province. Therefore, it is necessary to standardize the data of each index first. Since the indicators selected in this study are positive indicators, we can directly use the following formula to standardize the scores of each index:

For standardization, the index value before standardization, the maximum value of the index before standardization, the maximum value of the index before standardization, the minimum value of the index before standardization, the maximum value of the index after standardization, the minimum value of the index after standardization. In this study, after standardized processing, all index scores are projected to the [0.2,1] interval, so the value is 1 and the value is 0.2..

In this study, the Kmeans clustering algorithm is implemented in the Matlab environment. The standardized data is used as the input data of the Kmeans algorithm, and 17 sample points are input, and each sample contains 9 attribute elements. On this basis, we use Matlab's [Idx, C, sumD, D]=Kmeans (X, K) function to create a Kmeans algorithm model. Among them, X is a data matrix of N x P, K is a cluster number of cluster, Idx is a vector of N x 1, which stores a cluster number of each sample point, C is a K x P matrix, and the location of a K cluster center is stored, sumD is 1 x K and vectors, storing the sum of all the sample points between the classes and the clusters in the cluster. The distance between each sample point and all cluster centers. The cluster center number K is 4, and the clustering repeat number is 1000 times. The algorithm is used to get the result of the intensive use of cultivated land in Hubei province.

Using the Kmeans algorithm model constructed in this study, by setting a reasonable number of cluster centroid points, the cultivated land use intensive degree can be clustered in Hubei Province, and the result of the cultivated land use intensity zoning in Hubei province is obtained. The intensive degree of cultivated land utilization in Hubei province is divided into four levels, which are the western low intensive area, the eastern high intensive area, the southern medium intensive area and the central fine management area.

The low intensive areas in Western China include Shiyan, Enshi and Shenlong shelf forest in the west of Hubei province. In this area, the terrain is mainly mountainous, the terrain is undulating, the economic development is slow, the area of cultivated land is small and scattered, the intensity of cultivated land and the utilization intensity are low, and the level of intensive use of cultivated land is the lowest in Hubei province. Under the premise of protecting the ecological environment, we should make rational use of high-quality cultivated land resources and improve the utilization efficiency and output rate of cultivated land.

The highly concentrated eastern region is mainly concentrated in the northeastern part of Hubei Province, including Suizhou, Xiaogan, Wuhan, Huanggang and Ezhou. The region is the core of Wuhan city. It is the core area of Wuhan city circle and Hubei province. The input intensity of cultivated land is high, but the contradiction between the supply and demand of cultivated land is outstanding, so the intensive degree of the fine management area can not be reached. The region should strictly control the degree of "non farming" and "non grain" of cultivated land, and at the same time, improve the ability of high and stable grain production.   Moderate intensive areas in the South include Yichang and Jingzhou in the central and southern parts of Hubei, Xianning and Huangshi in the southeast. The area is flat and has the most abundant water and soil resources, and the Yangtze River has been through. However, the reasons why these cities are only in the intensive use level of low cultivated land are different. Jingzhou is the core area of Jianghan Plain, with the largest number of arable land in Hubei Province, but the low input intensity and utilization strength have become the constraints of the high intensive management of the cultivated land. Although the input intensity of Huangshi is not low, the output efficiency is not high. This area should improve the infrastructure of agricultural fields as soon as possible, and improve the scale,

intensification, industrialization and standardization of grain production.

The central fine management area is distributed from north to south in Hubei, Xiangyang, Jingmen, Xiantao, Qianjiang and Tianmen. As a traditional agricultural region, the region also has unique natural conditions and abundant soil and water, while the three cities of Xiantao, Qianjiang and Tianmen are located in the Wuhan city circle. While the intensity of input is guaranteed, the intensity and efficiency of the cultivated land are at a high level.

5. Conclusion and Discussion

Intensive use of cultivated land is an important way to ensure food security in the process of rapid urbanization. Studies have shown that due to regional differences in natural conditions and socio-economic development level, there are significant regional differences in cultivated land use intensification level. The zoning of cultivated land intensive utilization is the basis for realizing the division of cultivated land utilization. This study takes 17 cities and states in Hubei as an example, and constructs the index system from 3 aspects of input intensity, utilization intensity and output efficiency, and uses the Kmeans algorithm to divide the intensive degree of cultivated land. Kmeans algorithm is a common method in many clustering analysis algorithms. This method is simple and fast, and in the process of Kmeans clustering analysis, cluster core is used as the representative of the class, which can be given by the user and can be specified by the system. At the same time, for the same sample, different K values may result in different clustering results. Therefore, in application, cluster analysis should be combined with practical problems to select a reasonable K value.

The results showed that:

(1) The Kmeans algorithm was used to carry out the regionalization of cultivated land intensive utilization in Hubei province.

(2) There is obvious spatial difference in cultivated land intensive degree in Hubei province. In this paper, 17 cities and states in Hubei are divided into 4 regions: the western low intensive area, the eastern highly intensive area, the South moderate intensive area and the central fine management area. The intensive use of cultivated land in different cities in the same area is different. It is necessary to analyze the 3 aspects of the input intensity, the utilization intensity and the output efficiency of the cultivated land.

This study explores the application of Kmeans algorithm in the area of cultivated land use intensive degree, reveals the spatial pattern of intensive use of cultivated land in Hubei province and state in 2013, provides a scientific method for regional cultivated land use division management, and is of guiding significance to the policy formulation of intensive use of cultivated land. However, the intensive use level of cultivated land is dynamic, and the mechanism of the spatial and temporal differences of cultivated land intensive use is not clear. In the further research, we should study the dynamic zoning of cultivated land intensive utilization based on the Kmeans algorithm.

References

[1] Jiang C, Dong C, Bai-Hong L I. Research of Cultivated Land Management Zones Based on the Evaluation of Farmland Productivity——Taking Qingzhou City,Shandong Province as an Example[J]. Journal of Anhui Agricultural Sciences, 2016.

[2] Hao Z, Lei G, Yang X, et al. Utilization Zoning of Reserve Resources for Cultivated Land Based on Two-dimensional Graph Theory Clustering Method at County Scale[J]. Transactions of the

Chinese Society for Agricultural Machinery, 2017.

[3] Zhang H, Zhao X, Ouyang Z, et al. Spatial Disparity Features and Protection Zoning of Cultivated Land Quality Based on Spatial Autocorrelation——A Case Study of Shanggao County,Jiangxi Province[J]. Research of Soil & Water Conservation, 2018.

[4] Zhou M, Wang Z. Spatial Distribution of Primary Farmland Based on Cultivated Land Quality and Spatial Clustering at County Scale——A Case of Tuanfeng County,Hubei Province[J]. Research of Soil & Water Conservation, 2016.

[5] Yang J, Fan X U, Long Z, et al. Protection Zoning of Cultivated Land Comprehensive Quality Based on Spatial Autocorrelation [J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(3):156-163.

[6] Wang X, Zhao X, Guo X, et al. Protection zoning of cultivated land based on local spatial autocorrelation in Guixi City[J]. Acta Agriculturae Zhejiangensis, 2017.