



**HAL**  
open science

## Research on Evaluation of Soil Fertility Resource Space Based on Regional Hotspots and Clustering Method

Guifen Chen, Jian Lu, Ying Meng, Dongxue Wang

► **To cite this version:**

Guifen Chen, Jian Lu, Ying Meng, Dongxue Wang. Research on Evaluation of Soil Fertility Resource Space Based on Regional Hotspots and Clustering Method. 10th International Conference on Computer and Computing Technologies in Agriculture (CCTA), Oct 2016, Dongying, China. pp.246-256, 10.1007/978-3-030-06155-5\_24 . hal-02179961

**HAL Id: hal-02179961**

**<https://inria.hal.science/hal-02179961>**

Submitted on 11 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Research on Evaluation of Soil Fertility Resource Space Based on Regional Hotspots and Clustering

## Method

Guifen Chen<sup>1</sup>(✉), Jian Lu<sup>1</sup>, Ying Meng<sup>1</sup>, Dongxue Wang<sup>1</sup>

<sup>1</sup>School of Information Technology, Jilin Agricultural University, Changchun  
130118, Chinaguifchen@163.com

**Abstract:** Based on each unit of Nong'an County, With ArcGis software platform, the establishment of a part of the total nitrogen in the soil attribute data Nong'an topsoil fertility resources, phosphorus, potassium and organic matter and effective nutrient database, Global Clustering and outlier analysis and hot spot analysis of the autocorrelation method, the county topsoil fertility resources were analyzed and evaluated. The results for farmland protection and optimal layout policy to provide a reference for the theory and methods, but also on the implementation of the agricultural work provides precise basis for decision making.

**Keywords:** soil fertility resource; soil nutrient; regional hotspots and clustering method; GIS

## 1 Introduction

Topsoil soil resources refers to as agricultural production in a most basic means of production, cultivated land is the basic resources for human survival and development, and the natural elements in agricultural production has shown the potential production capacity, to promote farmland protection and scientific determination of cultivated land spatial distribution pattern and variation characteristics and other aspects has important significance. Topsoil soil resources

---

<sup>1</sup> Guifen Chen(1956-), College of Information Technology, Jilin Agricultural University, Changchun, Jilin, China

is a concept of cultivated land quality in the agricultural workers, soil fertility, crops grow well, high yield; poor soil fertility, is crop growth is not good, low yield. Criteria for evaluation of soil resources is the most direct should be the fertility of the soil, and measure indicators of soil fertility levels tend to use the organic matter content, the level of soil nutrient and soil physical properties. Therefore, the attribute data of soil nitrogen nutrient, available phosphorus, available potassium and organic matter were analyzed and researched, in order to provide the basis for the implementation of Xiang area tillage soil resource function zoning and differential control and precise farmland operation.

In the first law of geography, it is pointed out that all things are related, and that the things which are far from far away have higher relativity. In other words, the geographical data are related to each other due to the influence of spatial interaction and spatial diffusion. Space correlation analysis originated in the mid-19th century biometric study has now become one of the basic methods of Theoretical Geography, mainly used to analyze the spatial data statistical distribution. The analysis results are often dependent on the spatial distribution of data and different to traditional econometrics. The analysis method is realized by spatial autocorrelation index, this paper uses  $G_i^*$  Getis-Ord (hot analysis) and Moran's I (clustering and outlier analysis) to analyze the spatial autocorrelation of cultivated land quality

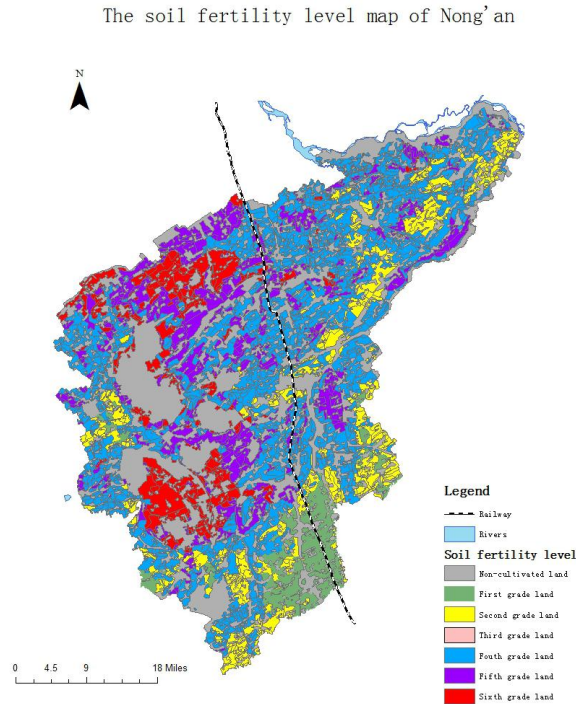
With the aid of GIS analysis technology, using the above two methods of spatial analysis, Nong'an County topsoil fertility resources part attribute data of soil total nitrogen, available phosphorus, available potassium and organic matter spatial hot spots and clustering and abnormal value are analyzed and provide significant test, the conclusion of the two methods for implementation of Xiang area tillage soil resource function zoning and differential control and precise farmland operation provides decision-making basis.

## **2 Materials and Methods**

### **2.1 The Basic Situation of the Study Area**

The Nongan County, Jilin Province, Nong'an county at Songliao plain, in the central part of Jilin Province, belonging to Changchun City, located 60 kilometers northwest of Changchun City, The latitude 43 degrees 54 ' to 44 degrees 56 ' , Longitude 124 degrees 32 ' to 125 degrees 45 ' , Second Songhua River, Yitong River, Xinkai River, Yinma River, passing through Weng Kehe. South and Changchun suburban neighborhood, Dehui City, bordering the East, northeast and Fuyu, across the river, north, Qian Gorlos Mongol Autonomous County, west of Changling, Gongzhuling City in the southwest and border. County 114.7 kilometers long from north to south, east-west width of 97.7 kilometers, the total area of 5400 square kilometers. In 2013, Nong'an County grain planting area

reached 366 thousand hectares, seed coverage reached more than 98%. The soil in Nongan county is divided into 10 soil groups, 20 subgroups, 50 soil genera and 111 species of soil, black soil, chernozem soil zonality. The local soil has meadow soil, alluvial soil, saline alkali soil, swamp soil and sand soil. According to the survey results of soil fertility level, the soil fertility level map of Nong'an GIS technology produced as shown in Fig.1 based on.



**Fig. 1.** The soil fertility level map of Nong'an

## 2.1 Research Methods

Spatial autocorrelation is a geographic feature attributes in a certain area of space characterization. With the help of GIS analysis techniques using Getis-Ord  $G_i^*$  (Hot Spot Analysis) and Moran's I (cluster analysis and outliers) in two spatial autocorrelation analysis of Nong'an part tith soil fertility resource attribute data TN ( N), phosphorus (P), available potassium (K) and organic matter of space clustering and hot spots and outliers are verified and the corresponding significance test, validation and test results show that: the two methods of analysis as a result, the job can accurately implement township gamut topsoil fertility resource function zoning and differential control farmland and provide a basis for decision making.

### 2.2.1 Getis-Ord $G_i^*$ (Hot Spot Analysis)

Hot spot analysis of each element of the data set can be calculated Getis-Ord  $G_i^*$  statistic (called G-i- asterisk). By z score and p-values obtained may know the location of clustering feature high or low values occur in space. Working way: view adjacent elements of the environment in every element. High value of the

element is often easy to attract attention, but may not have a statistically significant hot spots. To become a hot demand significant statistical significance, the elements should have a high value, and is surrounded by other factors also have a high value. Elements of a partial sum of its neighboring elements will be compared with the sum of all the elements; when there is a big difference when the sum of the partial sum of the partial expected and that they can not be the result of randomly generated, it will produce a significant statistical significance of the z-score. Its expression is as follows:

Getis-Ord local statistics can be expressed as:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{[n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2]}{n-1}}} \quad (1)$$

Wherein  $x_j$  is the attribute value of the element  $j$ ,  $w_{i,j}$  is the space between the right elements  $i$  and  $j$  weight,  $n$  is equal to the total number of elements, and:

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad (3)$$

$G_i^*$  is  $z$  score statistics, there is no need for further calculations.

For each feature dataset returned  $G_i^*$  statistic is a z-score. For positive  $z$  having a statistically significant score, the higher the  $z$  score, high values (hot spots) clustering more closely. For statistically significant negative  $z$  score on,  $z$  lower the score, the low (cold spots) clustering more closely.

### 2.2.2 Moran's I (clustering and outlier analysis)

Moran's simultaneously according to feature locations and feature values to measure spatial autocorrelation. Under a given set of elements and attributes related to the case, the expression pattern was clustering mode, the discrete mode or random mode. By computing Moran's I index value,  $z$  score and  $p$ -values for the significance index were evaluated.  $p$  value is based on the known distribution curve derived from the area of approximation (by the test statistic restrictions). Its expression is as follows:

Spatial autocorrelation Moran's I statistic can be expressed as:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{i,j} z_i z_j}{S_0 \sum_{i=1}^n z_i^2} \quad (1)$$

Wherein  $z_i$  is a property of its elements  $i$  mean  $(x_i - \bar{X})$  deviation,  $w_{i,j}$  is space right elements  $i$  and  $j$  weight,  $n$  is equal to the total number of elements,  $S_0$  is the weight of all polymeric space right:

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{i,j} \quad (2)$$

Statistics of  $z_I$  -score calculated according to the following form:

$$z_I = \frac{I - E[I]}{\sqrt{V[I]}} \quad (3)$$

Among them:

$$E[I] = -1/(n-1) \quad (4)$$

$$V[I] = E[I^2] - E[I]^2 \quad (5)$$

If the data set values tend to occur on spatial clustering (high values gathered in the vicinity of other high value; low gathered in the vicinity of other low value), the Moran's I index will be positive. If the high value of the exclusion of other high-value, low and tend to close, the index will be negative. If the positive and negative values of the cross product cross product offset value, the index will be close to zero. Since the molecules are normalized by the variance, so the value of the index will fall within the range of -1.0 to +1.0.

### 2.2.3 Spatial Autocorrelation of the Two Algorithms Similarities and Differences

Spatial autocorrelation statistic is a measure of geographic data (geographic data) of a fundamental nature: the degree of interdependence of the data in such a position with other positions between. Moran's I and Getis-Ord  $G_i^*$  is an important method of calculating spatial autocorrelation of two. These two approaches can be effective on the characteristics of spatial distribution of elements in space phenomena were analyzed. Getis-Ord  $G_i^*$  similarities and differences and Moran's I are as follows:

Algorithmically principle: common is the use of two algorithms are the average of all the properties of the overall assessment of the data analysis for identifying local spatial autocorrelation phenomenon. The difference is, Moran's I for lower weights

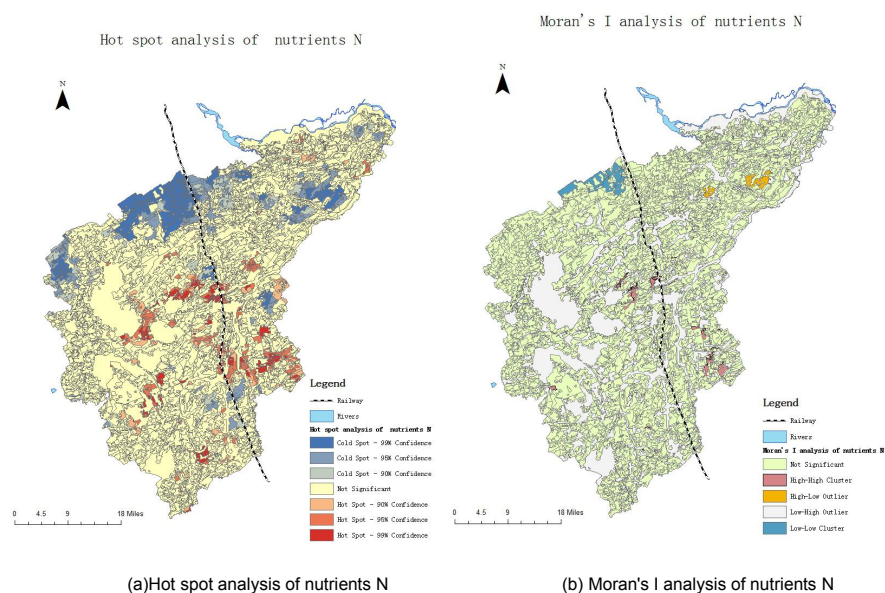
attention; 1 represents upright element having contain the same or similar low-high property values approaching element, the element is part of the cluster, negative 1 represents a near containing different values element, the element is an abnormal value, while Getis-Ord  $G_i^*$  weights sensitivity higher than the sensitivity of Moran's I; this algorithm is the sum total of the partial sum is compared with the sum of the partial when expected great differences that can not be randomly generated when the result is, will produce z having a significant score.

Results: the common denominator is two algorithms will compute p-values and z-score. The difference is, Moran's I would explicitly given the high value of the spatial concentration (HH), whispered spatial concentration value (LL), low high spatial concentration value (LH) and the level of spatial concentration value (HL) result Fourth, through this four things that can tell us whether there is a gathering space or outliers, and Getis-Ord  $G_i^*$  only compute p-values and z-score.

### 3 Results and Analysis

Use ArcGIS spatial analysis function, use the Hot Spot Analysis nutrient N (Getis-Ord  $G_i^*$ ) and Cluster and Outlier Analysis (Anselin Local Moran's I) tool Nong'an part tilth soil fertility resource attribute data, P, K and organic matter analysis and study, get a hot nutrient distribution Nong'an Nong'an charts and nutrient distribution Moran's I analysis chart.

#### 3.1 Hot spots, clustering and outlier analysis of nutrient N

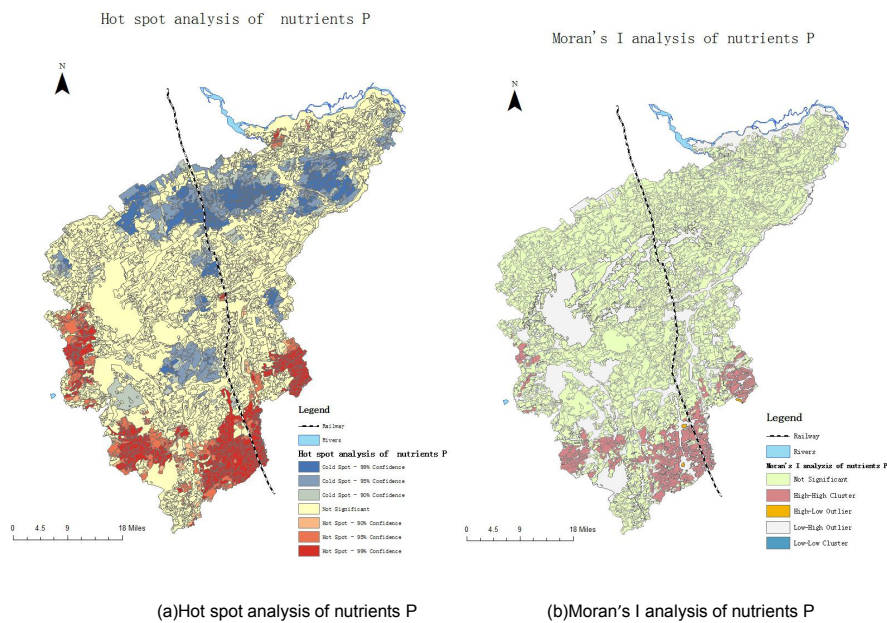


**Fig. 2.** Autocorrelation analysis space of nutrient N

Fig.2 (a) shows, the cold spot region nutrients N mainly Yang Shulin countryside near of Nong'an, hot spots mainly in the field of agricultural villages County

Central region; Fig. 2 (b) visible, whispered value appears in the spatial concentration Yang Shulin countryside near, township found in poplar scarce nutrient content of N, N nutrient content domain surrounding villages equally scarce, high value appears in the spatial concentration of agricultural communes County central region domain, it is found in the middle of nutrient rich Nong'an N content around N domain nutrient content villages equally rich, high and low values of spatial agglomeration occurs in XinNong countryside of Nong'an, it is found in the surrounding countryside, the new domain nutrient content of N is rural and less compared to the XinNong countryside.

### 3.2 Hot spots, clustering and outlier analysis of nutrient P

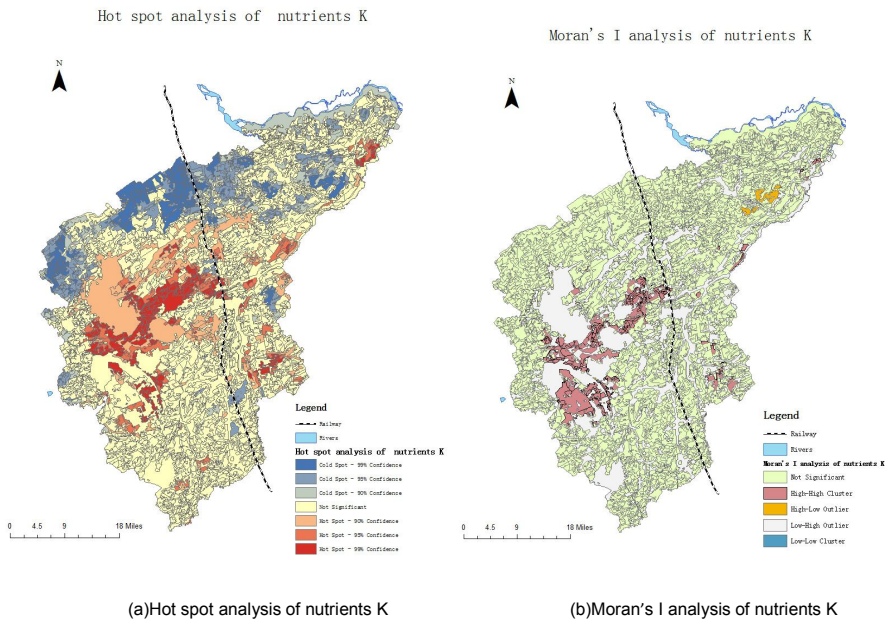


**Fig. 3.** Autocorrelation analysis space of nutrient P

From Fig. 3 (a) visible, P nutrient cold areas mainly distributed in Yang Shulin, Ha Lahai and Xin Nong countryside of Nong'an County, regional hot spots mainly distributed in Nong'an Fu Longquan ,Bao Jia, He Long countryside near; by Fig. 3 (b) shows that P Moran 's I analysis graph was mainly manifested as high value of spatial agglomeration, distribution situation and hotspot analysis of regional hot spots are generally the same, shows that P content rich in Fu Longquan , San Gang, Bao Jia of Nong'an country and around the township P equally rich.

### 3.3 Hot spots, clustering and outlier analysis of nutrient K

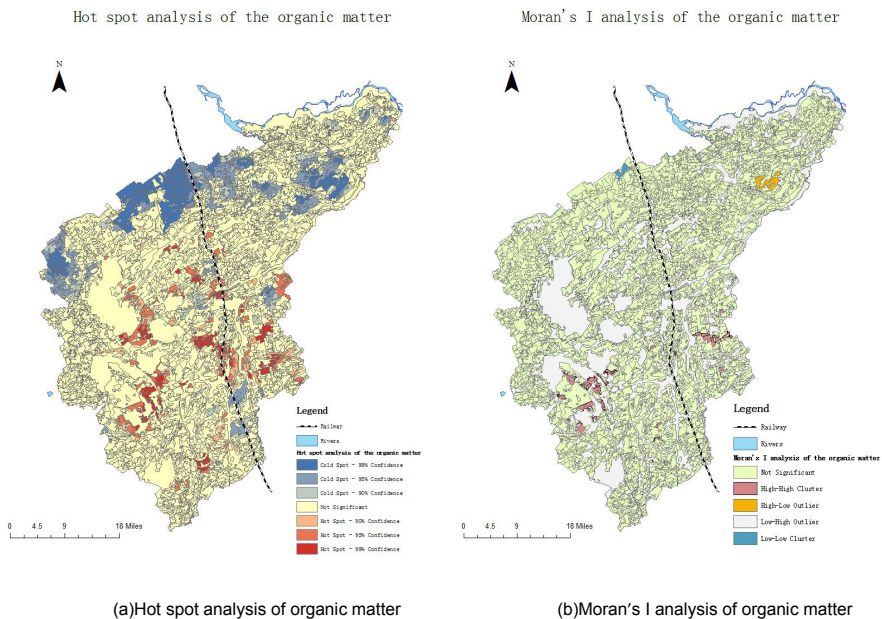




**Fig. 4.** Autocorrelation analysis space of nutrient K

Fig. 4 (a) visible, K cold areas mainly distributed in Yongan, Yang Shu Lin, Ha Lahai countryside of Nong'an County, regional hot spots mainly distributed in Nong'an County Central Rural Areas; by Fig. 4 (b) shows that high value space agglomeration appear in Nong'an County in Central rural area shows that in Nong'an County Central K content rich, around the country K content in the same rich, high and low values of the space agglomeration appear in Nong'an County Xin Nong countryside shows that in around the Xin Nong countryside domain K content and new agricultural township, compare less.

### 3.4 Hot spots, clustering and outlier analysis of organic matter



**Fig. 5.** Autocorrelation analysis space of organic matter

From Fig. 5 (a) visible, organic matter in cold areas mainly distributed in Yongan, Yang Shu Lin, Ha Lahai, Xin Nong countryside of Nong'an County, regional hot spots scattered distribution in Nong'an County Central Rural Areas; by Fig. 5 (b) shows that high value space agglomeration appear at the junction Ba Jilei and Long Wang countryside of Nong'an County shows that rich in organic matter content in the region, around the country organic matter content in the same rich, high and low values of the space agglomeration appear in Xin Nong countryside of Nong'an County, known in the content of organic matter in Xin Nong countryside around the township and new agricultural township, is less.

#### **4 Conclusions**

Through the analysis of the difference of each township Nong'an county domain resources space unit soil fertility, results show that:

- (1) Nong'an County topsoil fertility resource attribute of soil total nitrogen, available phosphorus, available potassium and organic matter and nutrient data has accumulated to a certain extent, based on Getis-Ord  $G_i^*$  on soil N, P, K and organic matter spatial hot spots were analyzed. Research and analysis shows that the N nutrient and organic matter regional hot spots distribution similar, hot spots are mainly distributed in Nong'an County in the south central part of the township area, and the freezing point of water is mainly distributed in Nong'an County, the northern part of the township area;
- (2) using Moran's I distinguish Nong'an County topsoil soil resources of soil property n, P, K and organic matter on nutrient data of space agglomeration and discrete spatial regions, the research result shows that the, different types of cultivated land quality index showed a strong spatial correlation, namely nutrient data high value or low value of Xiang area in space showed significant concentration of the state, and the aggregation degree is higher;
- (3) Getis-Ord  $G_i^*$  and Moran's I in the plough layer soil resource analysis based on, for the protection of arable land partitioning provides new ideas, the cultivated land quality in rural area of differential protection and fine management has a certain reference value.

#### **Acknowledgments**

This work was funded by the China Spark Program.2015GA66004. "Integration and demonstration of corn precise operation technology based on Internet of things"

## Reference

1. Ren Ping, Wu Tao, Zhou Jieming. Analysis of spatial distribution pattern and evolutionary characteristics of cultivated lands based on spatial autocorrelation model and GIS platform—A case study of Longquanyi District, Chengdu, China [J]. Chinese Journal of Eco-Agriculture.2016(03)
2. Hong Shuman, Hao Jinmin, Zhou Ning, Li Chen, Zhen Yu Lu. Change of cultivated land and its impact on grain production pattern in Huang-Huai-Hai Plain [J].Transactions of the Chinese Society of Agricultural Engineering.2014(21)
3. Li Yong, Zhou Yongzhang, Zhang Chengbo etc.. Application of Local Moran's I and GIS to Identify Hotspots of Ni,Cr of Vegetable Soils in High-Incidence Area of Liver Cancer from the Pearl River Delta,South China [J]. Environmental Science 2010 (06)
4. Xiong Chang Sheng, Shi Chuan Wei, Luan Jolin, Hu Yueming. Bspatial Pattern Disparity and Cultivated Land Quality Based on Moran's I Analysis[J].Resources science.2014 (10)
5. Zhao Jianjun, Zhang Hongyan, Wang Yeqiao etc. Research on the Quality Evaluation of Cultivated Land in Provincial Area Based on AHP and GIS:A Case Study in Jilin Province[J] Chinese Journal of Soil Science.2012 (01)
6. real Oman, Keung, Cheng Dao congruent. Analysis of the farmland productivity evaluation scale effect based on the grain yield level[J]. Journal of Henan Agricultural University. 2015 (01)
7. Feng Ting, Zhang Fengrong, Lican, Qu Yanbo, Zhu Fengkai. Spatial distribution of prime farmland based on cultivated land quality comprehensive evaluation at county scale [J]. Transactions of the Chinese Society of Agricultural Engineering. 2014 (01)
8. Shen Renfang, Chen Meijun, Kong Xiangbin, Li Yongtao, Yanan, Wang Jingkuan, Li Tao, Lu Mingxing. Conception and Evaluation of quality fo arable land and Strategies for its management [J]. Acta Pedologica Sinica. 2012 (06)
9. Zhang Yali, Zhou Yang, Zhai Xiaoyi, uo Shijie Wei Lin, Jun Chang Huang, Cheng Daoquan. Scale Transformation Method and Its Application of Cultivated Land Productivity Evaluation Between Count and City[J]. Journal of Soil and Water Conservation. 2015 (04)
10. Luo Jing, Chen Qiong Liu Fenggui, Zhang Yili, Zhou Qiang. Methods for reconstructing historical cropland spatial distribution of the Yellow River-Huangshui River valley in Tibetan Plateau[J]. Progress in Geography. 2015 (02)
11. Peng Meng, Hao Jinmin, Zhou Ning, Hong Shuman. Difference analysis of effect of rapid urbanization on cultivated land changes in Huang-Huai-Hai plain [J]. Transactions of the Chinese Society of Agricultural Engineering. 2013 (22)
12. Zhang Xiaoyu, Xie Xinpeng, Zhang Anlu. Spatial Unbalanced Development and Diffusion Route of Farmland Conversion in Wuhan City[J]. Journal of natural Resources. 2014 (10)
13. Chen Dan, Zhou Qigang, He Changhua, Liu Linlin.Analysis on Influencing factors of Cultivated Land Based on MPI in Typical Southwest Mountain Area--Taking Shizhu

- County of Chongqing as Example[J]. Research of Soil and Water Conservation. 2014(02)
14. Chen Shaokuan, Wei Wei, Mao Baohua, Guan Wei. Analysis on urban traffic status based on improved spatio-temporal Moran's I [J]. Acta Physica Sinica.2013 (14)
  15. Zhao Yeting, Chang Qingrui, Li Zhipeng, Liu Jiaqi. Spatial Characteristics of Soil Organic Matter and Total Nitrogen in Cultivated Land of Weibei Tableland Area[J]. 2014(08)
  16. long Jun, Zhang Liming, Shen Jinqian, Zhou Biqing, Mao Yanling, Qiu Longxia, Xing Shi and. Study on the spatial interpolation method of cultivated soil organic matter in the soil organic matter of the complex landform type area [J]. Journal of soil science.2014 (06)
  17. Zhang Chutian, Yang Yong, He Liyuan, Ji Shimin, Liu Yingying. Prediction of spatial Distribution of soil organic matter based on Environmental factors and a joint Probability Method [J]. Acta Pedologica Sinica.2014 (03)
  18. Qi Yanbing, Chang Qingrui, Liu Mengyun, Liu Jing, Chen Tao. County-Scale Spatial Variability of Soil Nutrient Distribution and Determination of Reasonable Sampling Density [J]. Chinese Journal of Soil Science. 2014 (03)
  19. Liu Zhipeng. Spatial Distribution of Soil Nutrients and the Impact Factors across the Loess Plateau of China [D]. The University of Chinese Academy of Sciences (Institute of Soil and Water Conservation ). 2013
  20. Wu Tao. Study on Spatial Evolution and Influencing Factors of Cultivated Land Based on GIS: A Case of Longquanyi District [D]. Sichuan Normal University 2015
  21. Zhao Yeting. Spatial Characteristics and changes of soil nutrients in cultivated land of Guanzhong region in Shaanxi Province based on GIS [D]. Northwest Agriculture and Forestry University 2015
  22. Zhu Jianfei. Research on the GIS — based Fertility Evaluation System of the Farmland in Nong'an County [D]. Jilin University 2008
  23. Tian Rui. Evaluation of Cultivated Land Fertility at Yongning County Based on GIS [D] Ningxia University 2013
  24. Tang beauty. GIS Supported Cultivated Land Fertility Analysis and soil Texting and fertilizer Recommendation at County Level--A case study in Guangrao County of Shandong [D]. Shandong Agricultural University, 2008
  25. Wan Limin. The evaluation of Land Productivity in Jilin Base on GIS [D]. Northeast Normal University 2014