

# DEFINITION OF MANAGEMENT ZONES OF SOIL NUTRIENTS BASED ON FCM ALGORITHM IN OASIS FIELD

Xin Lu<sup>1,\*</sup>, Yan Chen<sup>1</sup>

<sup>1</sup> Key laboratory of Oasis Ecology and Agriculture of Xinjiang Production and Construction Group, Shihezi, Xinjiang, P. R. China 832003

\* Corresponding author, Address: Key laboratory of Oasis Ecology and Agriculture of Xinjiang Production and Construction Group, Shihezi 832003, Xinjiang, P. R.China, Tel: +86-993-2057555, Email:lxshz@126.com

**Abstract:** The objective of this research was to define management zones of oasis cotton field. The variables of organic matter, available N, available P and available K data determined in 193 topsoil (0-30cm) samples were selected as data sources. Fuzzy c-means clustering algorithm was used to delineate management zones. In order to determine the optimum fuzzy control parameters, the fuzziness performance index (FPI), c-φ combinations and the multiple regression based on external variable were used in this study. Meanwhile, the cotton yield was chosen as the external variable. The whole field was divided in four management zones. And fuzziness exponent was 1.6. The zoning statistics showed that variation coefficient of soil nutrients decreased, while the means of the soil nutrients differed sharply between management zones. The average confusion index was 0.19 in all management zones. The overlapping of fuzzy classes at points was low and the spatial distribution of membership grades was unambiguous. The results indicated that fuzzy c-means clustering algorithm could be used to delineate management zones by selecting the appropriate external variables. The defined management zones can be used for fertilizer recommendation to manage soil nutrient more efficiently.

**Key words:** Management zones; Fuzzy clustering; Soil Nutrients, external variables, Oasis cotton field

## 1. INTRODUCTION

Study on the precise management zone in the basis of multi-type soil characters in the inner-land oasis, not only can master the oasis agriculture ecological environment feature and the special production function, but also is very important for reasonable development and sustainable development of the oasis agriculture. With regards to the soil characters variance, the management that cut the whole area into the different mean texture zones was become a focus spot in the precise agriculture (Schepers et al, 2004). The fuzzy *c*-mean algorithm (FCM) was most widely used in the management zones. But FCM method exists fuzzy control parameter (fuzzy type number and fuzzy weighting exponent) can not be easily determine; the analyze ability is worse for the factors in the different kinds of boundaries (Shi Hongbo et al, 2003). They have been become the continuous explore theme. De Bruin (1998) stated the multi-fitting methods on the basis of extraneous vary. The method didn't obtain the information from the algorithm method, but determine the reasonable fuzzy control parameter by analysis of the extraneous vary and spatial distribution relation of subordination degree for fuzzy data collection. The paper researches the big farmland in the inner-land oasis agriculture mode, use the multi-kind of soil character as the data resource, and apply the fuzzy *c*-mean classification method to define the field management zones. On the basis of inform study, making the cotton yield as the extraneous vary. At the same time adopted the no-determine complex degree of the fuzzy subordination relationship with confusion index (CI) surface sample to assessment to the feasible for the management zone for the big field in the northern oasis agriculture mode.

## 2. MATERIAL AND METHOD

### 2.1 Research Area Survey

The experiment region was No.148 region of Xinjiang production construction army-group. It lies in the 44°46'50"-44°59'25"N, 86°9'10"-86°22'51"E. The climate belongs to the continental climate, the rainfall is rare, and the resources of light and heat are very rich. The topography is smooth, besides the fluctuation partially; there are mainly gray desert soil, Aeolian sandy soil and salt soil, in the total the five second-type and nine soil families. The crop in the region mainly plant cotton, the cultivated area is 12,800ha, the agriculture management lacks of central management, the quantity fertilizer including extra fertilizer is the same in the block.

## 2.2 Soil data obtain and sample analysis

In October ,2006, we collected the soil sample according to the no-regular way in the 148 region drip irrigation zone, the sample centralization on the each sample spot was the centralization of circular, radius was 4.5m. at the depth of 0-30cm in the soil main cultivate in the range of circular collected five soil sample and combined, the sum of soil sample was 193. At the same time, measure the yield in the range that was length of 5.86m, width of 2.25m, is the centralization on the basis of the soil sample. The field of cotton is the output in the ha unit area. The soil sample collected became dry in the air and need to sieve out the stones from the soil, and then measured in the lab. The projects of each soil sample include the OM, AN, AP, AK and total salt. (Bao Shidan, 2000).

## 2.3 Fuzzy classification and interpolation

Fuzzy *c*-mean algorithm is a no-surveillance algorithm method, which is widely used in the classification of soil, topography, terrain features, output and remote data. For multi-resource of a spatial sample number *n*, soil character number *p*, fuzzy *c*-mean algorithm can determine the reasonable fuzzy type number *c*, define the algorithm centre and computer the fuzzy subordination for each classification class.

The basic idea is make iterated objective function become minimized; the function was used as following:

$$J(MC) = \sum_{i=1}^n \sum_{j=1}^c \mu_{ij}^{\varphi} d_{ij}^2 \quad i = 1, \dots, n; j = 1, \dots, c \quad (1)$$

In the middle of equation (1), *C* is algorithm type centroid matrix *c*×*p*, *M* is sample fuzzy subordination degree matrix *n*×*c*,  $\mu_{ij}(\mu_{ij} \in [0,1])$  is the *j* algorithm type subordination degree number to the soil character for the *i* sample spot,  $\varphi(\varphi \geq 1)$  is a control algorithm result and type overlapping degree constant factor(Gao Xinbo et al, 2000), also is the fuzzy lighting exponent,  $d_{ij}^2$  is the distance square between the *i* soil sample spot character and the *j* type centre.

Outside standard analysis is built on the basis of extraneous vary; which is not use the algorithm analysis, but is impacted by the outside condition impacted the algorithm. For the outside standard analysis, if the stronger relationship between the spatial variance of extraneous varies and the spatial distribution of spatial subordination degree of soil nutrient, the fuzzy *c*-mean algorithm can better express the effect of the soil data algorithm. The multi-fitting method built by de Bruin and Stein computed the relationship

between fuzzy subordination degree and extraneous vary. The fitting model of fuzzy subordination degree and extraneous vary shows in the equation (2). In the equation (2),  $\beta_0$ 、 $\beta_1$ 、... $\beta_c$  are the fitting coefficients,  $\varepsilon$  is random error, the check decision coefficient  $r_a^2$  in the fitting model can express the relation between the different c value and  $\varphi$ , when  $r_a^2$  is in the peak, the c value and  $\varphi$  is most best reasonable. In the middle of equation (3),  $r^2$  is the decision coefficient in the fitting equation,  $n_s$  is the sample number of extraneous vary.

$$\text{Extraneous vary} = \beta_0 + \beta_1 \mu_{A1}(x) + \dots + \beta_c \mu_{Ac}(x) + \varepsilon \quad (2)$$

$$r_a^2 = r^2 - \frac{c(1-r^2)}{n_s - c - 1} \quad (3)$$

In the study, extraneous vary is the cotton yield data in 2006 year according to the result as following.

In the basis of algorithm class result, make the partially subordination relationship of the different soil nutrient content for the sample soil normal kriging method and get the sole category fuzzy chart for soil nutrient. At the same time, in order to determine the overlapping degree between the sample spot for the different soil character type, use the confusion index to estimate the fuzzy no-determine for the soil sample subordination relationship (Burrough et al, 1997).

$$CI = 1 - [\mu_{\max i} - \mu_{(\max-1)i}] \quad (4)$$

In the equation (4),  $\mu_{\max i}$  is the biggest subordination degree value for the certain soil type character for the  $i$  soil sample spot,  $\mu_{(\min-1)i}$  is the second subordination degree for the soil subordination relationship.

In the study, Fuzzy c-mean algorithm is successfully used by software MATLAB7.0; the variance feature analysis of fuzzy subordination degree, semi-variance function computing, theory model fitting, kriging method and chart drawing are both completely made by the software of GS+3.5 and ArcGIS9.2.

### 3. RESULTS AND ANALYSIS

#### 3.1 Descriptive statistics results

Descriptive statistics of soil nutrient and cotton yield in the research region (Tab.1), recognition check of data in the data value domain does not

exist the outlier. From the mean value, according to the Xinjiang soil survey management (Xinjiang Production and Construction Corps Soil Survey Office, 1993), in the research region, the content of organize material (OM) and available phosphorus belongs to the middle level, the content of available nitrogen is lower than middle level, the content of available potassium is richer. According to the vary number CV, the variance degree can be roughly estimated, the results are made as following: the weak variance of CV is smaller than 10%; middle variance of CV is lie in the range of 10 ~ 100 % ; Strong variance of CV is bigger 100 % (Wang Shaoqiang et al, 2001), in the research zone, soil nutrient was impacted by large-scale mechanization of agriculture work and personal factor under membrane drip for long years, the middle variance is still showed the soil spatial character in the zones and can satisfied with agriculture management zone divided and the prerequisite condition to vary fertilized management.

Tab.1: Descriptive statistics of soil properties and cotton yield

Soil properties	Min	Max	Mean	SD	CV(%)
OM(g/kg)	3.80	22.50	10.37	2.92	28.18
AN(mg/kg)	20.00	102.00	47.63	12.85	26.98
AP(mg/kg)	3.60	40.60	12.29	5.66	46.03
AK(mg/kg)	161.00	519.33	318.02	58.26	18.32
Cotton yield(kg/ha)	1297.42	2925.15	1970.18	374.49	19.00

### 3.2 Fuzzy classification result output of soil nutrient

Bring the soil nutrient fuzzy collection  $n \times p$  ( $n=193$ ,  $p=4$ ) into fuzzy c-mean method and make fuzzy algorithm on the basis of the related method above.

Adopted multi-fitting method on basis of extraneous varies to determine the reasonable fuzzy control parameter. The result of method permits simultaneously choice  $c$  and  $\varphi$ , it is same with  $c-\varphi$  multi-combination. Take the cotton yield dada in 2006 in the research region extraneous vary to compute the fitting model between the yield and the fuzzy subordination degree of the different  $c$  and  $\varphi$ , the change relationship between the adjusted coefficient of model and  $\varphi$  shows in Fig 1. We can get the result,  $r_a^2$  is more and more when  $\varphi$  is more and more,  $r_a^2$  reaches peak value 0.57 when  $c=4$ ,  $\varphi=1.6$ .

Fuzzy control parameter determined are  $c=4$ ,  $\varphi=1.6$ , the whole region is divided into four soil nutrient management zones, class centroids of soil properties and Partial memberships and confusion index of the sampled soils to the classes shows in the Tab.2 and Tab.3. In order to examine the overlapping degree of subordination degree, the CI of each sample also shows

Tab.3. Class centroids that that is the average of the character vary for multi-dimensional spatial can reveals the variance of soil character data of the different sample and data combinations (Tan Manzhi, 2006). According to the soil type central value data in the Tab 2, make an appraisal for the soil management zones, the result shows that the content change of soil nutrient character is stability. In the management zone No.1, the content of OM and AN is highest, the nutrient content is richest; on the contrary, the management zone 2 the whole nutrient is lowest; the content of management zones 3anf 4 are in the middle of the range. The output result of fuzzy algorithm method contacts the character spatial and geography spatial distribution information of soil multi-nutrient, it is very important for the variable fertilizer applied, making the management measure in management zones.

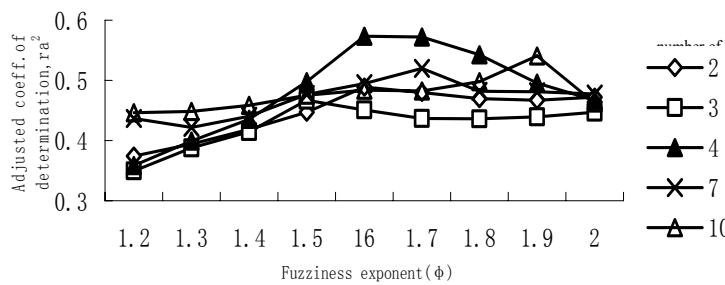


Fig.1: Adjusted coefficient of determination as a function of class partial memberships for different combination of  $c$  and  $\phi$

Tab.2: Class centroids of soil properties

Fuzzy class	OM(g/kg)	AN(mg/kg)	AP(mg/kg)	AK(mg/kg)
C1	13.19	50.62	15.71	411.15
C2	8.19	42.52	10.04	222.89
C3	10.89	47.31	12.48	345.27
C4	9.95	48.18	11.96	293.50

Tab.3: Partial memberships and confusion index of the sampled soils to the classes

Soil sample	C1	C2	C3	C4	CI
1	0.001	0.028	0.014	0.957	0.071
2	0.086	0.022	0.625	0.267	0.643
⋮					
193	0.001	0.944	0.004	0.052	0.108

### 3.3 Visual express of class result

The spatial information contacted in the output result of fuzzy algorithm method only can as the base of spatial forthcoming and can not directly show

the important act of continuous class in the spatial forthcoming (Tan Manzhi, 2006). So, we need to take the geography-statistics method complete the visual express of the quota subordination relation on the geography spatial for the different soil sample, directly reveal the feature of spatial distribution and variance of soil nutrient on the whole research region, furthermore estimate the soil fertilizer of the research region. On the base of data on the Tab.3, obtain the semi-variance function of subordination degree, and then make the normal kriging method for the partial subordination degree of each soil sample, at last get the subordination degree of the different fuzzy type and reality the spatial forthcoming of soil fertilizer distribution in the research region.

Tab.4: Semivariogram parameters of fuzzy classes

Zones	Model	C <sub>0</sub> Nugget	C+C <sub>0</sub> Sill	C <sub>0</sub> /(C+ C <sub>0</sub> )/%	Major Range(km)	Major Range(km)
Zones 1	Guassian model	0.37	2.75	13.46	5.57	2.88
Zones 2	Exponential model	0.56	2.44	23.14	4.35	2.23
Zones 3	Spherical model	0.34	2.93	11.60	6.81	3.47
Zones 4	Spherical model	0.36	2.56	13.87	5.10	2.92

The variance function theory model and their parameter shows in Tab.4. The model of management zone in the No1, 2, 3, 4 is Guassian model, Exponential model, Spherical model, Spherical model respectively. The main axle is change in the range of 4.35~6.81km; Second axle range is change in the range of 2.23~3.47km. The ratio values of C<sub>0</sub>/(C+ C<sub>0</sub>)/% is lower than 25%, which shows the stronger spatial auto-relation (Chang et al, 1998).

Judged from the Fig.4, the subordination degree of the northern soil for the management zone2 is higher than else region. It is near the Gurbantonggut desert and exist irrigation fixed Aeolian sandy soil family and irrigation soda saline soil Aeolian sandy soil family. There are a great deal of new-cultivated field, the plant age is shorter; the cultivated layer is fuzzy; the ability of conserve water and fertilizer is worse. The soil fertilize is poorer. The subordination degree of the middle soil for the management zone4 is highest (Fig.2d), the region is the distribution of discard yield and main exist the discard saline soil family. The depth of underwater is higher; the water hardness degree is bigger; the content of soil layer is higher; the nature plant is rare; the soil lateral nutrient is lower. the subordination degree of the northern-west and southern-west soil for the management zone3 is higher(Fig.2c), the region main exist the irrigation soda saline gray desert soil family, irrigation SO<sub>4</sub><sup>2-</sup>cl<sup>-</sup> saline gray desert soil family and irrigation cl<sup>-</sup>SO<sub>4</sub><sup>2-</sup> saline gray desert soil family and exist the certainty second saline. The soil nutrient content shows the slowly increasing from the west to east and from the southern to northern, the whole fertility level is between the

management zone4 and the management zone1. The subordination degree of the southern soil in the middle part for the management zone1 is higher than 0.8 (Fig.2a). It means the soil nutrient content is highest. It is reason for the distribution the yellow and red texture second type of irrigation-cultivated gray desert soil and sand-gravel texture of irrigation-cultivated gray desert, the ability of conserve of water and fertilizer is stronger than else region, the cultivated age is longer, the degree of cure is higher, the soil nutrient content in the cultivated layer is higher. The oasis management distribution chart on the basis of the subordination degree shows in the Fig.3, which is play a key act for the sole fuzzy subordination degree spatial distribution of soil nutrient.

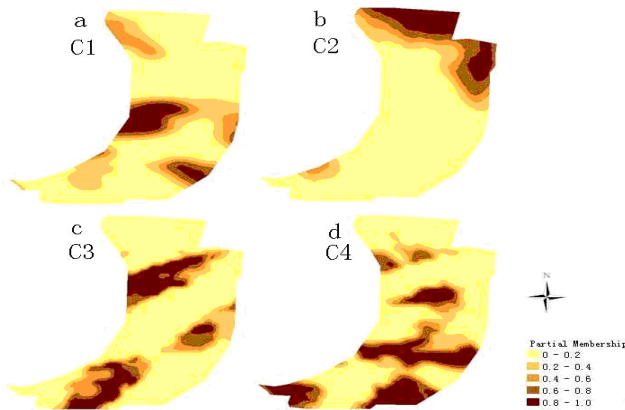


Fig.2: Predictive maps showing spatial variation of partial memberships of soil to individual fuzzy classes derived respectively

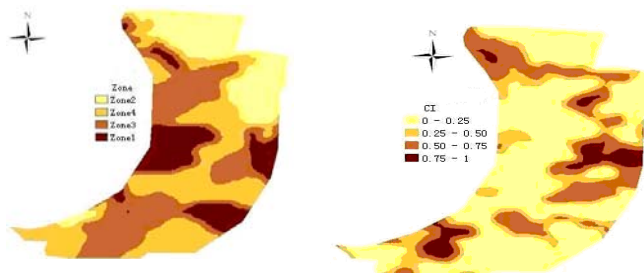


Fig.3: Management zones maps of oasis Fig.4: Predictive maps showing spatial variation of confusion index field

The confusion index means the no-determine complexion of sample spot fuzzy subordination relationship. The value is bigger, the overlapping degree of the different fuzzy type in the sample soil is more higher, no-determine of soil subordination for the sample soil is more higher; on the contrary, it means the lower over-lapping degree, the soil subordination relation is slowly explicit. When  $CI=0$ , It means that the character spatial and



geography spatial of sample soil exists the explicit lateral. The confusion index picture is made as following in the Fig.4, from the picture, we can see that the confusion index in the across of new-cultivated, discard and water resource et al and main cultivated field is relative higher. In a certainty zone in the middle and southern in the study zone, CI value is highest,  $CI > 0.8$ , it shows that in these zones the no-determine of the subordination relationship of soil character content is bigger. The mean value CI of soil subordination relationship is 0.19 in the zone, in the whole view, the no-determine for the fuzzy class output result of soil character data is more smaller, the soil subordination relationship in the spatial geography is relative explicit.

### **3.4 Check of management zones**

In order to determine if management zone is the accordance of vary fertilization, make the normal analysis for the soil character data in the management zone and use LSR method to check the variance sufficient (Tab.4), compared with no divided zones, the vary value of soil character in the each management zone all is reduce. The vary value of OM in the whole zone is reduce from 28.18% to the range of 18.94%~25.57%, the vary AN is reduce from 26.98% to the range of 18.81%~25.15%, the vary value of AP in the whole zone is reduce from 46.03% to the range of 31.46%~42.28%, the vary value of total salt in the whole zone is reduce from 18.32% to the range of 1.53%~13.55%. These are show that the variance degrees all are reduce in the management zones, the content distribution becomes average.

The check of variance sufficient show: judged from 0.05 remarkable level, the mean value of OM and AK both reach sufficient variance in the different zones; the management zones of AN do not reach the sufficient variance between 1 and 3 do not exist; the management zones of AP also do not reach the variance between 2 and 4, between 3 and 4, for the two type nutrient among of else management zones exist the variance. Judged from the 0.01 remarkable level, the management zone of OM between 3 and 4 and the management zone of AN between 1 and 3 do not exist the deficient variance, for the two type nutrient among of else management zones exist the efficient variance; the management zones of AP among of 2, 3 and 4 can not reach the efficient variance, but management zone among of else management zones exist sufficient variance; In the whole view, the spatial distribution variance of soil character in the management zone is more smaller after divided, but the variance among of management zones exist variance. From the pre-text judge that bring the same amount fertilization in the

management inner, but in the different management zones take the different fertilization.

#### **4. DISCUSSIONS**

For the fuzzy c-mean algorithm, it needs to determine the management  $c$  and fuzzy index  $\phi$ ; The paper take the multi-fitting method on the basis of extraneous vary to determine the reasonable management zone number and fuzzy index for the different fuzzy type number on the basis of two indexes. The method can better reveal the relationship between the variance of extraneous vary and spatial distribution of subordination degree for fuzzy data collection. The whole zone can provide four management zones, fuzzy lighting exponent is 1.6. the experiment shows that variance value of soil character is reduce compared with the management zones, soil character of management zones is variance efficient, at the same time in the whole zone confusion index is more lower, the overlapping degree for the different fuzzy type is more smaller, the soil subordination relationship in the geography spatial is relative precise. It proves that the measure that management zone is used in the large-scale farmland is feasible. The result of management zone can efficiently express the spatial distribution information of soil character and provide the policy basis for the precise and high efficient field management, at the same time, provided the theory basis in order to study the precise agriculture management.

#### **ACKNOWLEDGEMENTS**

This study was supported by the fifth national technology supported plan of P. R China (grant no. 2006BAD21B02-2) and the Program for Xinjiang Production and Construction Group (grant no.IRT0412).

#### **REFERENCES**

- Gorsevskil P M, Gessler P E, Jankowski P. Integrating a fuzzy k-means classification and a Bayesian approach for spatial prediction of landslide hazard. *J Geograph Syst*, 2003, (5): 223-251.
- Lai Xianqi. *China oasis agronomy*. Beijing: China Agricultural Press, 2005 (in Chinese)
- Li Xiang, Pan Yuchun, Zhao Chunjiang, Wang Jihua, Bao Yansong, Liu Liangyun, Wang Jindi. Delineating precision agriculture management zones based on spatial contiguous clustering algorithm. *Transactions of the CSAE*, 2005, 21(8): 78-82 (in Chinese)

- Li Y, Shi Z, LI F, Li H Y. Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land. *Computers and Electronics in Agriculture*, 2007 (56): 174–186.
- Li Yan, Shi Zhou, Wu Cifang , LI Feng, Cheng Jieliang. Definition of management zones Based on fuzzy clustering analysis in coastal saline land. *Scientia Agricultura Sinica*, 2007, 40(1):114-122 (in Chinese)
- Schepers A R, Shanahan J F, Liebig M K, Schepers J S, Johnson S H, Jr A L. Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. *Agronomy Journal*, 2004, 96:195-203.
- Shi Hongbo, Yu Jian, Huang Houkan. An effective implementation method of Fuzzy c-Means algorithm. *Journal of the China railway society*, 2003, 25(1):63-67 (in Chinese)
- Zhang Fenghua, Lai Xianqi. Oasis agriculture characters and development in arid area of northwestern China. *Journal of Arid Land Resources and Environment*, 2003, 17(4):19-23 (in Chinese)