

The Application of the OPTICS Algorithm in the Maize Precise Fertilization Decision-Making

Guowei Wang^{1,2,3(✉)}, Yu Chen^{1,3}, Jian Li^{1,3}, and Yunpeng Hao^{1,3}

¹ School of Information Technology, Jilin Agricultural University,
Changchun 130118, China

41422306@qq.com, 123929697@qq.com,
2312852319@qq.com, 793097534@qq.com

² School of Biological and Agricultural Engineering,
Jilin University, Changchun 130022, China

³ Jilin Province Research Center,
Changchun 130118, China

Abstract. With the development of computer science and information technology, data mining technology in the field of agriculture in recent years has become a hot research. Corn planting process, rational fertilization can effectively promote the growth of corn, however, no basis and targeted fertilization may cause shortage of low fertility soil fertilization, high fertility soil fertilization overdose. To solve this problem, In this paper, cluster analysis OPTICS algorithm based on density of soil classification, and press the nutrient balance method to calculate the level of soil fertility for each corresponding amount of fertilizer, farmers can be targeted based on fertilizer fertilization. In the town of Yushu City, Jilin Province by Gongpeng for application, compared with the traditional fertilization, fertilizer input savings of 20.5 %, maize yield of about 10 %, not only to meet the needs of farmers, but also achieve a reduction in fertilizer inputs, increase production purposes.

Keywords: Data mining · Precision fertilization · Fertility classification · OPTICS algorithm

1 Introduction

Precision fertilization is one of the core elements of precision agriculture [1]. The method using GPS positioning to obtain field information, including production monitoring, soil sampling and so on. The computer system through analysis of the data processing, decision of the management measures of agricultural land, the yield and soil status information is loaded with GPS devices, implemented on each operating unit due to the overall balance of the soil due to crop fertilization, thus greatly improving fertilization fertilizer use efficiency and economic benefits, reduce the adverse impact

on the environment [2, 3]. However, the variable rate fertilization machine with GPS has not been used widely, so the method of dividing the operation unit is not conducive to the promotion of farmers.

Soil fertility is an important aspect of the constitution of land productivity, it depends on many factors and a variety of nutrients in the soil organic matter content, texture, thickness and soil tith configuration, and these factors on soil fertility and sizes mutual restraint, is the ability to provide the soil for plant growth and coordination of nutrition and environmental conditions. The soil fertility grading, contribute to the establishment of scientific fertilization scheme, abandon the traditional concept of the more, the better. The soil nutrient status, crops need fertilizer characteristics determine the reasonable amount of fertilizer and nutrients mix, thereby reducing the production chain of agricultural input costs and the probability of the occurrence of pests and diseases, improve nutrient utilization. Reduce the background of cultivated land, control the point source pollution and non-point source pollution, and restore the ecological agriculture and sustainable development [4–7].

Currently, soil fertility studies, Luo Laijun [8] used fuzzy mathematics theory, to explore the reclamation area of farmland soil fertility grading method, which provides a reasonable reference for the target yield. Based on GIS Map info software, the establishment of land resources management information spatial database and attribute database. The soil nutrient content and distribution map of soil nutrient content and distribution map were made.

In summary, this paper adopts cluster analysis in optics algorithm based on density and on soil fertility grading, and according to the nutrient balance method to calculate the corresponding of soil fertility at each level of fertilization, provides a feasible scheme for large area promotion of precision fertilization.

2 OPTICS Algorithm

2.1 Basic Concepts of OPTICS Algorithm

The core idea of the density of clusters is a point of ϵ neighborhood neighbor points to measure the density of the point where the space [9]. If ϵ neighborhood neighbor exceeds a specified threshold MinPts, it is that the point is in a cluster, called the core point, or that the point is on the boundary of a cluster, called a boundary point [10, 11]. Here are some definitions:

Define 1. Core distance

The core distance object p is p as the core object of minimum. If p is not the core object, the core distance of the p is not meaningful. The core distance is the smallest radius of the neighborhood of the core points.

To explain the definition of core distance and reachable distance, we set up Minpts = 5, $\epsilon = 6$ mm, if the core object, a field point p must have at least 5 points (including itself). In Fig. 1, the d point is the boundary point that satisfies the core object of p , so the core distance of p is $c = 3$ mm.

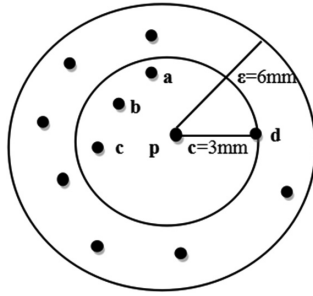


Fig. 1. Core distance

Define 2. Reachability-distance

Reachability-distance is the bigger one from the core of p and the Euclidean distance between n and p.

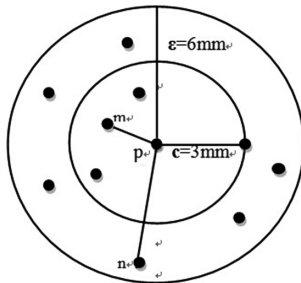


Fig. 2. Reachability-distance

In $Minpts = 5$, $\epsilon = 6$ mm, according to n p on the definition of distance, distance of n to Euclidean distance p. Because the Euclidean distance of p to n is more than the core distance of 3 mm.

Define 3. Directly density-reachable

If the p is the core point, q in the ϵ -neighborhood of p, the p direct density of up to q.

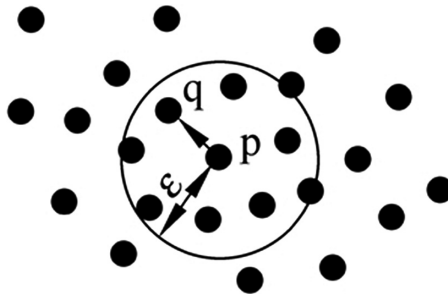


Fig. 3. Directly density-reachable

2.2 OPTICS Algorithm Process

The algorithm is as follows:

- Step1: Create two queues, order queues and result queues. (Orderly queue is used to store the core object and its core direct access to an object, and arranged in ascending distance; Results the sequence of the output sequence of the sample points is stored)
- Step2: If all the points in the sample set D are processed, the algorithm ends. Otherwise, choose an untreated (is not in the queue) and as the core object of sample points, find the all density reachable sample points, such as the sample point does not exist in the queue, will it into the ordered queue, and are ordered by distances of up to
- Step3: If the ordered queue is empty, the jump to the step 2, otherwise, the first sample point from the orderly queue to expand, and the sample points will be removed to the results queue
 - Step3.1: Judge whether the expansion point is the core object, if not, back to step 3, otherwise find the expansion point of all the direct density of up to point;
 - Step3.2: Judging whether the direct density of the sample point is already there is the result queue, is not processed, otherwise the next step;
 - Step3.3: If orderly queue already exists the directly density reachable point, if this new reachability distance less than old reachable from, sharp distance is used to replace old reachability distance, orderly queue reordering.
 - Step3.4: If the direct density of the sample points is not present in the ordered queue, then the insertion point is inserted, and the ordered queue is re ordered;
- Step4: The algorithm is over and the ordered sample points of the output queue are ordered.

3 Core Precision Fertilization Decision

3.1 Data Acquisition

Gongpeng town in Yushu No. thirteen villages, land use GPS to obtain information, press 40 m * 40 m mesh distance. Each grid soil sampling, 117 soil samples were obtained, each soil sample for testing to give each sample soil organic matter, available phosphorus, available nitrogen, available potassium content data.

3.2 Data Standardization

Since the practical problems, different data generally has a different dimension, in order to make different dimensionless quantity can be compared, so the need for standardization of data, about to compress the data to the [0,1]. Data normalization formula is:

$$g = \frac{\max - \min}{G - \min} \quad (1)$$

3.3 Soil Fertility Grading

Using OPTICS algorithm to process the data, that is, soil fertility, soil fertility of the clustering, divided into different clusters. This paper divides the soil fertility into four clusters. Calculate the distance of the range of each cluster, and draw up the distance chart of the sample points according to the reachable distance. The output sequence is the horizontal coordinates, and the distance chart is the vertical coordinate. The reachable distance ordered graph is generated by the plot function in the MATLAB software.

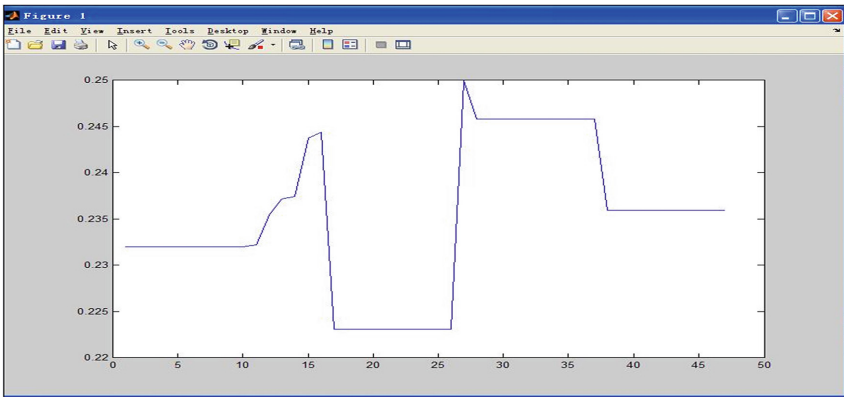


Fig. 4. Orderly figure of the reachability-distance

As shown in Fig. 4, it is clearly divided into four wave values, that is, the four clusters, we can according to the distance ordered map of the horizontal coordinates, determine the number of the four clusters contained in each of the soil number.

The point is, we can through the analysis of ordered graphs can be obtained directly when the parameter E (field radius) was 0.24, minpts (epsilon minimum points) for 10 of the clustering results. Can be seen from the chart, 0.24 the horizontal lines on the ordinate, data is obviously divided into three clusters, other not in the range of points is regarded as outlier.

According to the OPTICS algorithm, the soil fertility is effectively clustered into clusters. The average values of the four attributes of each cluster are respectively summation, and then according to the size of the classification, and the largest of the land, and so on.

According to the average values of the attributes of the four clusters, the results of the soil classification are shown in Table 1.

Table 1. Soil classification result

Soil level	Included sample points
Class A	a68 a102 a111 a75 a103 a92 a112 a105 a106 a96 a73 a72 a113 a95 a93 a108 a90 a79 a99 a109 a104 a85 a100 a81
Class B	a5 a83 a22 a60 a29 a15 a6 a97 a56 a37 a7 a17 a49 a8 a12 a26 a84 a59 a34 a16 a25 a23 a13 a24 a101 a38 a19 a116 a74 a45 a71 a94 a2 a18 a31 a3 a27 a47 a117 a20
Class C	a48 a44 a76 a11 a21 a30 a61 a4 a70 a33 a65 a50 a82 a14 a10 a36 a69 a51 a43
Class D	a41 a86 a78 a91 a28 a80 a115 a89 a88 a77 a39 a40 a42 a67 a55 a64 a46 a58 a35 a54 a110

3.4 Calculation of Fertilizer Application

Using nutrient balance method, we calculated the amount of soil nutrient, to meet the crop yield. Because the data in the nutrient balance method need to be the raw data of the sampling point, the data value is shown in Table 2.

Table 2. Average properties of different soil classification

Soil level	Organic mean	Available phosphorus mean	Available nitrogen mean	K mean
Class A	27.16404	11.24264	151.8825	82.20833
Class B	27.7042	15.93813	132.363	105.4524
Class C	30.28655	11.58181	128.2779	108.5789
Class D	26.6535	13.8507	129.7659	103.7143

Calculation of nutrient balance fertilization model by fertilizer application rate (2):

$$sf = \frac{cl \times xs - cd \times ys}{hl \times ly} \tag{2}$$

- sf: Fertilization;cl: Corn yield target;
- xs: Grain corn nutrient absorption amount per 100 kg;
- cd: Soil nutrient determination;
- ys: Soil available nutrient conversion factor;
- hl: Fertilizer nutrient content;
- ly: Fertilizer utilization season

Taking the mathematical model of fertilizer application rate in the elm city as an example (The content of P2O5 was 46 % in the application of diammonium phosphate):

$$\text{Phosphatefertilizer} = \frac{\text{Corn target amount} * 0.07 - 0.03 * \text{Soil nutrient content} * \text{Soil available nutrient conversion factor}}{0.46 * \text{Fertilizer utilization rate}}$$

$$\text{Soil available nutrient conversion factor} = \begin{cases} \frac{1578.8 * \text{Soil nutrient content}^{-0.98}}{100} \\ \frac{1068 * \text{Soil nutrient content}^{-0.832}}{100} \\ \frac{732 * \text{Soil nutrient content}^{-0.749}}{100} \end{cases}$$

$$\text{Blank area yield} = \frac{0.3 * \text{Soil nutrient content} * \text{Soil available nutrient conversion factor}}{0.022}$$

$$\text{Fertilizer utilization rate} = \begin{cases} \frac{(43.4 - 0.024) * \text{Blank area yield}}{100} \\ \frac{(36.6 - 0.025) * \text{Blank area yield}}{100} \\ (4.6 - 0.035) * \text{Blank area yield} \end{cases}$$

Yield = 10000 kg/hm²

According to the principle of nutrient balance, the Table 2 data and substitute it into formula (2), can be obtained as shown in Table 3 shows each level of soil fertility of specific fertilizer.

Table 3. Soil nutrient fertilizer rate

Soil level	N amount of fertilizer (kg/hm ²)	P amount of fertilizer (kg/hm ²)	K amount of fertilizer (kg/hm ²)
Class A	222.9694	195.0158	165
Class B	235.7228	166.1162	115
Class C	239.792	192.369	115
Class D	238.2537	177.1553	125

4 Conclusions

The application of OPTICS algorithm in precision fertilization has not been reported. The method in this paper, for example application in arch shed town of Yushu City in Jilin Province, the average fertilization amount was 332 kg/hm², compared with the traditional fertilization, and fertilizer 68 kg/hm²; the average yield of 8313 kg/hm² was 813 kg/hm², and the yield was increased. Indeed achieve a reduction in fertilizer inputs, improve soil environment, increase production, income purposes.

References

1. Helong, Yu., Guifen, C., Chunguang, B.: Corn precision fertilization Database Modeling. *Maize Sci.* **16**(4), 184–188 (2008)
2. Guifen, C., Li, Ma., Hang, C.: Research status and development trend of precision fertilization technology. *J. Jilin Agric. Univ.* **35**(3), 253–259 (2013)
3. Xiaohua, Q., Yilong, Z., Ronggen, H.: Precision fertilization technology and application measures. *Anhui Agric. Sci. Bull.* **7**(2), 44–45 (2001)
4. Xu, J.Y., Wbster, R.: Optimal estimation of soil survey data by geostatistical method-semi variogram and block Kriging estimation of topsoil nitrogen of Zhangwu country. *Acta Pedol. Sci.* **162**(4), 291–298 (1983)
5. Yost, R.S., Uehara, G., Fox, R.L.: Geostatistical analyst of soil chemical properties of large land areas iv. Semivariograms. *Soil Sci. Soc.* **46**, 1028–1032 (2000)
6. Mei, L., Xuelei, Z.: GIS-based evaluation of farm land soil fertility and its relationships with soil profile configuration pattern. *J. Appl. Ecol.* **22**(1), 129–136 (2011)
7. Yuhong, W., Xiaohong, T., Yanan, T.: Assessment of integrated soil fertility index based on principal components analysis. *J. Ecol.* **29**(1), 173–180 (2010)
8. Lajun, L.: Study on partition fertilization and soil testing and soil fertility of cultivated land reclamation area classification. *Mod. Agric. Sci. Technol.* **42**(2), 228–229 (2013)
9. Hongbo, Z.X., Shuo, B.: OPTICS-Plus for Text Clustering. *Chin. J. Inf.* **22**(1), 22–24 (2008)
10. Xiujie, W., Quanchao, Z., Haijun, L.: The value of clustering analysis and principal component analysis in the study of Anthropology. *J. Anthropol.* **29**(4), 35–37 (2007)
11. Shanjie, W.: Further thinking about the classification method. *J. North China Inst. Sci. Technol.* **5**(1), 108–110 (2008)