

An Application of Fuzzy C-Means Based Clustering Technique in Smart Farming

Tian zhuo¹, Libaicheng²

Information teaching and Management Center, Jilin Agricultural University, Changchun, Jilin, China
xiaomarkshaoqing@163.com, lbc@163.com

Keywords: Precision farming; Fuzzy c-means clustering; apparent electrical conductivity

Abstract. Customised farming has seen ever increasing interest of next level farming practices across the globe by the modern agronomist of the age. Customized Management practices (CMP) make use of creating management zones (MZs'), leading to precision farming through application of variable rate technology for soil inputs. In this paper site-specific Apparent Electrical conductivity (ECa) and crop yield information have been used to study the spatial distribution and variability of the land. Matlab 8.5 version was used to implement Fuzzy c-means clustering in order to cluster the data and various performance indices such as NCE, FPI and their inference were calculated to determine the clustering performance. Both the clusters (ECa and crop yield) were mapped and good amount of overlap was observed indicating the relationship between ECa and crop yield.

Introduction

Agriculture production system is an outcome of complex interaction of seed, soil, water and agro-chemicals. Therefore, judicious management of all the inputs is essential for the sustainability of such a complex system. It is also certain that availability of labour is also going to be a major worry for the agriculture lands. So it is imperative to utilize the resources judiciously and effectively. Precision farming merges the new technology borne out of information age and mature agriculture industry. It is an integrated crop management system that attempts to match the kind and amount of inputs with the actual crop needs for small areas in the field. This goal although is not new but the modern technology and advancements allow this concept to take shape. New information tools not only serve to manage within-field variability but also general aspects of farming such as documentation of field work. Precision agriculture (PA) emerged with the application of new technologies in farming such as GPS, new sensing techniques and variable rate technology, enabling local field measurement and adjustment of local inputs. But the precision agriculture system is not yet complete. Especially, decision support^[1] in precision management needs to be further developed so that detailed field information may be used in an effective way in precise field management.

The objective of the research described here is clustering of soil parameters, and comparing this with the yield results. Another clustering was performed, using apparent soil Electrical Conductivity (ECa) and yield of Punjab Agriculture University (PAU) fields. This was used as a means to highlight the soil-yield interaction in the field. The interpolation methods and grid size of the soil and crop data can influence the analysis result.

Materials and Methodology

Survey data of geo-referenced ECa of soil has been collected on three experimental fields.

Sites description. Sensor measured soil properties and yield were measured during the 2010 season, on field scale basis with summer paddy, varieties such as PBW 343, PBW 550 and DBW 17. The normal crop protection measures were taken, including herbicides, stem growth reduction and fungicides. The field was harvested in May 010. Numbered plots with dimensions as per below mentioned details were established on the field. Post harvesting different varieties of paddy fields^[2] during 2010 summers, the field survey was conducted on these three experimental fields situated in

Punjab agricultural University (PAU), Ludhiana, (INDIA).

All the fields were under experimental precision farming studies, located at farm and machinery site of PAU. An arrangement of field1 was divided into three different parts (part1, part2 and part3) each having gridmatrix structure of 3(column) by 5(row's) total 15 grid cells in each part, each grid cell was having dimension of 6m by 5m, a row of one metre width was separating two parts as shown in Fig.1. The sensor data was collected at the vertices of each cell. Therefore, twenty four points have been collected from each matrix grid i.e. seventy two points from three parts in total of the field1 with the help of in-situ field usable sensors in one case, whereas collected fine resolution geo-referenced continuous ECa measurements with geonics EM38MK2 trolley mounted arrangement from the same field.

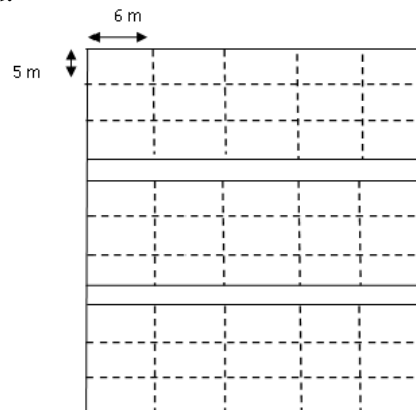


Figure 1. Arrangement of experimental field

Field2 as shown in Fig.2 was organized in to 4 by 5 grid matrix, each grid cell having dimensions of 9m by 9m giving thirty geo-referenced point's whose measurements were done at the vertices of each grid cell besides recording continuous geo-referenced ECa measurements.

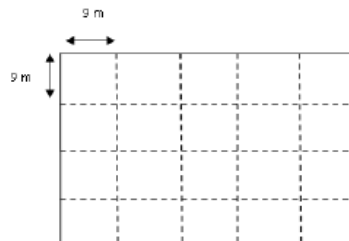


Figure 2. Grid arrangement of field 2

Similarly, as shown in Fig.3, field3 was divided into 1by1 grid matrix, each grid cell^[3] having dimension of 9m by 9m creating 24 geo-referenced points at the vertices of grid cells.

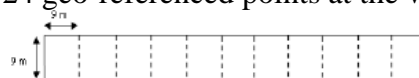


Figure 3. Grid arrangement of field 3

Figure 3. Grid arrangement of field 3

Data Collection of geo-referenced ECa. The basic steps followed for the data collection of geo-referenced ECa during survey are as follows: Firstly spatial variations of ECa data was collected with EM38MK2 field usable geo-sensor, data was logged manually as well as automatically. The data logging system consists of geonics EM38MK2xp logging software, laptop as a field computer, and associated cables to connect the computer to the geonics EM38-MK2 conductivity sensor. Through EM38MK2xp program, survey data was acquired and recorded from the EM38-MK2 system to field computer under the control of operator. using geo-sensors.

Data Analysis of Measured Soil Properties

Clustering. Clustering^[4] is the process of partitioning or grouping a given set of patterns into disjoint clusters. This is done such that patterns in the same cluster are alike and patterns belonging to two different clusters are different.

Fuzzy- C- Means Clustering. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters.

This method, developed by Dunn in [7] and improved by Bezdek in [8], is frequently used in pattern recognition. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2 \quad 1 \leq m < \infty$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the i th of d -dimensional measured data, c_j is the d -dimension centre of the cluster, and $\|*\|$ is any norm expressing the similarity between any measured data and the center.

Results and Discussion

The data was divided into two groups: one with the information about parameters: temperature, moisture, conductivity, pH, real dielectric and apparent electrical conductivity (Eca) and other with the crop yield along with their longitude and latitude information.

Both the data groups were separately clustered with fuzzy-c means algorithm. Optimum number of clusters was calculated using fuzzy performance index (FPI) and normalized classification entropy (CE).

The following figure shows the FPI and CE indices and hence clearly shows the optimum number of clusters as 3 for both yield and conductivity.

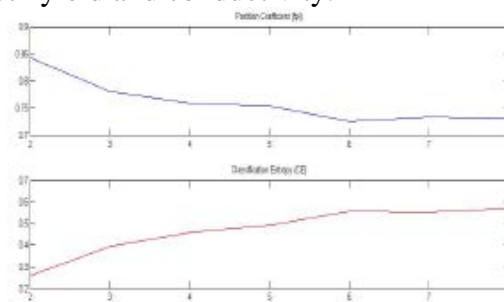


Figure 4. FPI and CE indices of conductivity

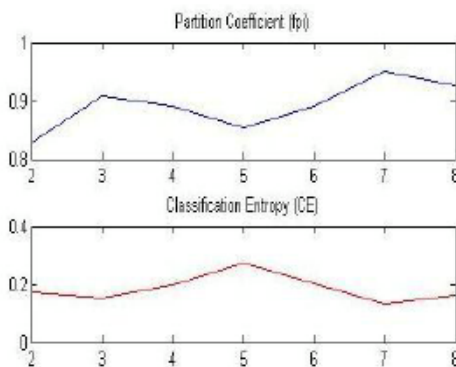


Figure 5. FPI and CE indices of Yield

Conclusion

Recent research in the area of precision agriculture has focused on the use of management zones as a more economical approach to define regions within a farm field for variable application of soil inputs rates. In the present work, the spatial variations of these limiting-yield factors were quantified using sensing techniques in terms of representative ECa and further processed with mathematical tool and aggregated into management zones by fuzzy c-means clustering algorithm. The resulting classified management zones using ECa can be used to characterize spatial variability in soil chemical properties^[5] and showed favorable agreement with crop yield variability pattern.

These results provide a basis of information for managing land in a rationally site-specific and precise way and may be used to develop a targeted soil sampling plan to capture variability in various soil properties that are likely to influence crop yield. Knowledge of such management zones may enable a reduction in the number of soil analysis needed to create application maps for certain cultivation operations. Since the cost of obtaining soil samples to characterize field variability is a key problem in precision agriculture, this is particularly advantageous. Based on these findings, it can be concluded that ECa and yield measurements can be reliably used for creating productivity zones on claypan soil fields.

References

- [1] Pierce F.J. and P. Nowak, "Aspects of precision agriculture," *Advances in Agronomy*, Vol. 67, pp. 1-85, 1999.
- [2] P. Robert, "Precision agriculture: research needs and status in the USA," *Proceedings of the Second European Conference on Precision Agriculture*, Odense, Denmark. Vol. 1, pp. 19-33, 11-15 July 1999.
- [3] M. Moore, "An investigation into the accuracy of yield maps and their subsequent use in crop management. Diss. Cranfield, Cranfield University, Silsoe College, Department of Agriculture and Biosystems Engineering. 1998.
- [4] Taylor R. K., G. J. Kluitenberg, M.D. Schrock, N. Zhang, J.P. Schmidt and J. L. Havlin, "Using yield monitor data to determine spatial crop production potential," *Trans ASAE*, Vol. 44, No. 6, pp.1409-1414, 2001.
- [5] Godwin R. J., G. A. Wood, J. C. Taylor, S. M. Knight, and J. P. Welsh, "Precision farming of cereal crops: a review of a six year experiment to develop management guidelines," *Biosystems Engineering*, Vol. 84, No. 4, pp. 375-391, 2003.