INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
12–062/ZIP/2012/14–4–481–491
http://www.fspublishers.org

Full Length Article

Analysis and Delineation of Spatial Variability using Geo-sensed Apparent Electrical Conductivity and Clustering Techniques

BABANKUMAR S. BANSOD[1], O.P. PANDEY† AND N.L. RAJESH‡
Agrionics (V-1), CSIR-Central Scientific Instruments Organisation, Sector 30C, Chandigarh-160030, India
†School of Physics and Material Science (SPMS), Thapar University, Patiala, India
‡Department of Soil Science, University of Agricultural Science (UAS), Chandigarh-160030, India
[1]Corresponding author's e-mails: scientist_babankumar@csio.res.in; baban79@yahoo.com

ABSTRACT

The present study was carried out to assess inter and intra field spatial variability based on Apparent Electrical Conductivity (EC_a) data at different depths of soil. Further, Principal Component Analysis (PCA) and Cluster Analysis (CA) were carried out to study the distinct soil variations based on field-scale EC_a measurements. PCA results of score plot were observed to verify pattern matching between measured soil spatial variability representatives i.e., EC_a, crop yield and variable input nutrient rates. The paper further reports delineation of management zones (MZs) using EC_a and crop yield data and PCA, hierarchical clustering and fuzzy c-means (FCM) clustering algorithms besides finding predictions for un-sampled spatial surfaces using univariate geo-statistical technique. Moreover, for determining optimal number of zones, clustering performance was measured using Fuzzy Performance Index (FPI) and Normalized Classification Entropy (NCE) indices. The results revealed that soil EC_a, nutrient rate and crop yield information could be quantified and aggregated using CA that characterize spatial variability among soil and crop productivity. © 2012 Friends Science Publishers

Key Words: Customized management zones; Principal component analysis; Spatial variability; Crop yield modelling; Fuzzy C-Means; Fuzzy performance index and Normalized classification entropy

INTRODUCTION

Precision Farming (PF) have proven track in the development of new technologies that have potential to increase farm yield at reduced agricultural inputs and environmental damage. Recent research in PF has focused on use of MZs as a method for variable application of inputs. MZs represent a homogeneous combination of potential productivity-limiting factors, which are therefore permanent (Fridgen *et al*., 2000a) and refer to geographic regions that present topography, crop and soil attributes with minimal heterogeneity (Luchiari Jr. *et al*., 2000). The determination of homogeneous areas within a field is difficult to achieve due to the complex combination among factors which may influence yield (Jose & de Cesar, 2008). Clustering soil and yield data give a starting point to analyse the causes of yield variability and can be a basis for defining MZs (Vrindts *et al*., 2003). Lot of research recently has been directed and evolving towards evaluating different techniques, algorithms, procedures, use of data layers, use of significant variables, for identifying precise MZs. This includes: use of single or multiple data layers such as EC_a, crop yield, topographic features (slope & elevation), nutrient levels, and remotely sensed images, data analysis procedures and methods includes: Fuzzy c-means (FCM), spatially contiguous K-means (SCKM), Geo-spatial and statistical techniques, kriged soil test point data, consideration of farmers knowledge, coefficient of variation of data layer, rasch model and PCA (Fridgen *et al*., 2000b; Moral *et al*., 2011). These computational techniques are used either alone or in combination. Further, since maps of soil physical properties (clay content), cation exchange capacity (CEC), organic matter, and yield show visible correlation with EC_a, use of EC_a as an indirect measure of soil physico-chemical properties is considered as a rapid and inexpensive tool for PF (Williams & Hoey, 1987; McBride *et al*., 1990; McNeill, 1992; Jaynes *et al*., 1994; Rhoades *et al*., 1999). These correlated properties have a significant effect on water and nutrient-holding capacity, drivers of crop yield, and hence can be used for delineating productivity zones for claypan soils (Jaynes *et al*., 1995; Kitchen *et al*., 2005). Determining the proper amount of nitrogen (N) to be applied to an agricultural field is a source of debate and discussion among growers, input suppliers, and researchers due to its serious environmental and economic concerns. With blanket N-rate there are typical areas within a field, which are over- or under-dosed. The usefulness of soil EC in PF is because it correlates well with soil texture: sands have a low conductivity, silts have a medium conductivity and clays have a high conductivity

(Lund *et al.*, 2001), and influences N movement through it. Lund *et al.* (2001) have presented three case studies to determine the proper amount of nitrogen to be applied to an agricultural field by assessing N-availability and its efficiency using EC_a measurement. Laboratory studies in the earlier reported works have shown that soil nutrients directly affect the electrical conductivity of the isolated soil solution (Ouyang *et al*., 1998) therefore, it is reasonable to assume that EC_a could be used to measure the available nutrient content of the soil, eliminating the need for time-consuming and expensive soil sample acquisition and analysis (Heiniger *et al*., 2003).

The present investigation deals with the assessment of relation between soil EC_a with crop yield response to interpret spatial variability. Further to delineate intra filed zones based on the soil EC_a, which would facilitate site-specific Customised Management Practices (CMP) enabling variable rate of soil input application. The Unscrambler software version 10.0 was used to investigate the relationship between conductivity and yield data. The interpretation of the field-scale measured EC_a data have been performed by PCA based scatter plots and hierarchical cluster (HCA) and FCM analysis. The specific objectives of this study were to: (i) identify variable N-rate and EC_a relation (ii) investigate the specific relationship between *yield spatial variability* and *EC_a spatial variability* (iii) Assess cluster performance by computing FPI and NCE to validate optimal number of zones.

MATERIALS AND METHODS

Modern on-the-go mapping technology of field-scale EC_a information was combined with traditional data collection (soil sample analysis & crop yield observations). The combination of EC_a, applied nutrients, soil and crop yield information was expected to give better insight to spatial and temporal variation in the field, leading to design site-specific input management scheme of the field.

Field survey and data collection: DGPS based survey data of soil EC_a were collected from three experimental fields, which were monitored under PF studies, located at Punjab Agricultural University (PAU), Ludhiana, INDIA. No-tillage system was practiced under paddy cultivation, during the summer cropping seasons and harvested in the month of May 2010. The soil type in this area is classified under tropical arid brown soil. A systematic study using field-scale soil EC_a measurements was conducted to assess the soil quality and from the site-specific crop management perspective, which includes following steps: (i) Establishing site/field boundaries, (ii) recording site metadata, (iii) collecting Trimble DGPS coordinate system, (iv) establishing EC_a measurement intensity, (v) geo-referencing site boundaries and significant physical geographic features with DGPS, (vi) measuring EC_a after completing necessary calibration and compensation process of geonics EM38MK2 field-usable geo-sensor, (vii) designing of sampling strategy based on geo-referenced EC_a data, (viii) soil sampling at specified sites designated by the sample design, (ix) lab analysis for the determination of soil properties, and (x) spatial and geo-spatial analysis to interpret EC_a with respect to crop yield and soil properties.

Sensor measured soil properties and manually measured yield data were collected from field 1, on field scale basis prior to harvesting. The fields were cultivated with paddy varieties such as V1=PBW 343, V2=PBW 550 and V3= DBW 17. During normal crop cycle, nitrogen fertiliser was applied in variable rates: N1, N2, N3, N4 and N5, with average dose of 0, 75, 125, 175, and 225 kg nitrogen/ha. The normal crop protection measures were taken, including herbicides and fungicides. The field scale data collection was done using geonics EM38MK2 conductivity meter and geo-referenced with Trimble DGPS. EM38MK2 is a portable device weighing about 3 kg and was pulled through the field by the operator using wooden trolley. The sensor was positioned at about three inches above the ground surface providing continuous EC_a data logging. It comprises of receiver and transmitter coil placed apart and electrically insulated from each other, transmitter coil induces changing field and produces eddy currents in the soil and secondary coil in turn generates "EMF" proportional to soil conductivity caused by soil eddy currents. EM38MK2 measures simultaneously both the quad-phase (apparent conductivity) and in-phase (susceptibility) components, within two distinct depth ranges, to a maximum effective depth of 1.5 m. It consists of two receiver coils, each in the vertical dipole orientation, separated by 1 m and 0.5 m from transmitter, simultaneously providing data at two depths, providing both shallow (0-0.75 m) and deep readings (0-1.5 m). This portable meter combines performance of all previous EM38 models and has features such as Bluetooth for wireless data transmission, extended battery life, automated field scale calibration etc., which provides more soil information in less time.

On-the-go, geo-referenced EC_a data for each field were collected using specially developed trolley free from any EM interferences since it was built from wood, a low induction number material. Other data logger system component consisting of DGPS were attached to this wooden trolley set-up for real time geo-referencing of EC_a measurements. The variance observed in the field-scale measured data for soil spatial variability is given in Table I.

Experimental field organization: Field1 was divided into three different parts (part1, part2 & part3) each having grid matrix structure of 3(columns) by 5(rows) totalling 15 grid cells in each part and each grid cell had dimensions of 6 m X 5 m. Field2 was divided into 4 by 5 grids, each cell having dimensions of 9 m X 9 m creating 30 points for measurement. Field3 was divided into 11 by 1 grids, each cell having dimensions of 9 m X 9 m creating 24 points for measurement. Experimental field organization for field 1, 2 and 3 are shown as below [Fig. 1 (a), (b) and (c)].

Soil sampling scheme using ESAP: After collection of field scale data, EC_a directed soil sampling was done using ESAP-RSSD (Response Surface Sampling Design software), a multi-program, statistical software package designed and distributed by the Salinity Laboratory, USA for optimizing: the sampling location and assessment, and prediction of soil salinity and related variables. Using this software, signal de-correlation was performed over the data; outliers were seen and removed by running signal validation module. After signal validation, configuration of data was further accepted based on RSSD algorithm; sampling design was created with different adjustment factors for the given samples to meet the optimality criteria (< 1.3). Resultant sample design footprint obtained was saved in .txt and .jpg format. Finally, optimal EC_a directed soil sample design was produced for each field. Twelve sampling points for each part of field1 and twenty sampling points for field2 were designed.

Soil sample collection: The sampling locations obtained from ESAP software were converted to DGPS compatible file so that the locations could be traced in the field. The sample design footprints obtained from ESAP were converted to DGPS waypoint file using DGPS pathfinder software. Also, the waypoint file was sent to DGPS device using data transfer option of the DGPS pathfinder software. After transferring waypoint file to DGPS, field was navigated and marked for different sampling locations within the field. At the designated sites, soil samples at 1feet and 2 feet depth were collected using manual core sampler. The soil samples were packed in air tight plastic bags and were labelled according to geo-referenced site details. Finally, soil samples were dried under shade, before sending to soil testing laboratory of PAU, Ludhiana (India) to determine their physico-chemical properties such as moisture, pH, real dielectric (Electrical Conductivity), Organic matter, clay and temperature. The data were divided into two groups: one with the information about parameters such as temperature, moisture, pH, real dielectric and apparent electrical conductivity (EC_a) and second group with the crop yield along with their longitude and latitude information.

Data analysis: After data collection, statistical and geo-statistical analysis was carried out to find correlations between EC_a, crop yield, N-rate and physico-chemical soil attributes. It includes: regression analysis, scatter plot analysis between EC_a and yield to check pattern matching, variability analysis using PCA and EC_a and yield data to find significant Principle Components (PCs) representing spatial variations and then applying FCM on significant PCs in MATLAB to delineate MZs. Finally to visualize spatial variations, kriging interpolation was applied by establishing appropriate theoretical semi-variogram models, firstly, for known sampled data and then for un-sampled locations. Statistical analysis software, UNSCRAMBLER 10.0 version was used to carry out PCA, HCA and correlation studies. MATLAB software tool was used to perform classification as well as validation using FCM technique and FPI performance index respectively. ArcGIS 9.0 was used to display MZ's. For post processing operation of DGPS data and sampling location tracing, Path Finder Office 4.10 program was used.

RESULTS AND DISCUSSION

PCA and cluster analysis of EC_a, crop yield and variable soil N-rate: The PCA scores plot was studied for EC_a, Crop yield, and variable soil N-rate. The scores plot is a map of samples, which shows how they are distributed. It can be used to isolate samples that are similar, or dissimilar to each other. The following inferences were drawn in the present study.

Inference 1: By using information about yield, EC_a and variable N-rate applied in different parts of field, the following observations were derived (Fig. 2a & b):

a) It is evident from PCA score plot that the properties of part1 and part3 show greater similarities as compared to that of part 2 since they lie close to each other in scores plot.

b) It can also be studied from the correlation loading plot (Fig. 2b) that EC_a measurements at two different inter-coil spacing configuration [vertical mode (CV) *i.e.* CV-0.5 m & CV-1.0 m] represents different regions of subsoil spatial variability. This may be of relevance, while understanding soil profile at different crop root zones, while deciding type of crop production.

c) Part2 showed the highest yield, and EC_a at 75 cm depth range (CV-0.5 m) has also shown increment in same direction. This is because values obtained for 0-75 cm depth range represent the region, which is mainly dominated by clayee sand, where EM fields are confined to show good conductivity in this region.

d) In field1 grain yield was significantly linear with applied variable nutrient rate until it reached its threshold value of 'N5' rate.

Above inferences were supported by PCA plot studies between geonics measured field scale variations represented by EC_a and the yield information (Fig. 2a & b).

Inference 2: The PCA plot between the nitrogen rates applied and crop yield information for field1, helped in finding threshold values of nitrogen to be applied within individual parts of field1to ascertain maximum yield. This level of interpretation, with the currently available information, further enabled monitoring of variable nitrogen rate control for a specific crop and field location. Below are the point-wise observations noted for individual parts of field1 derived from spatial EC_a variations through PCA analysis considering nutrients and crop yield variability.

a) The crop yield was found to be increasing with the amount of nitrogen (Fig. 3a & b) in part1, but on increasing nutrient-rate beyond a specific threshold (N4), the crop yield has shown downward trend in this part. That means when the supply and the requirement match, any further increase in N-rate does not account for crop

Fig. 1: Arrangement of experimental field 1, 2 and 3

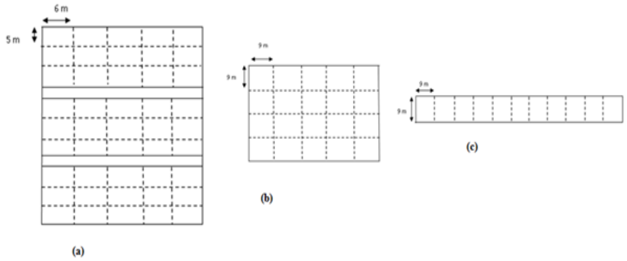
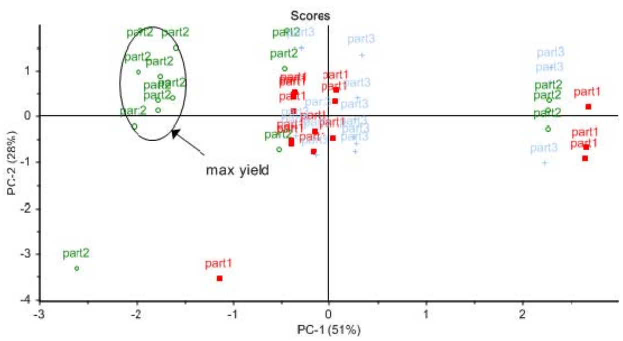
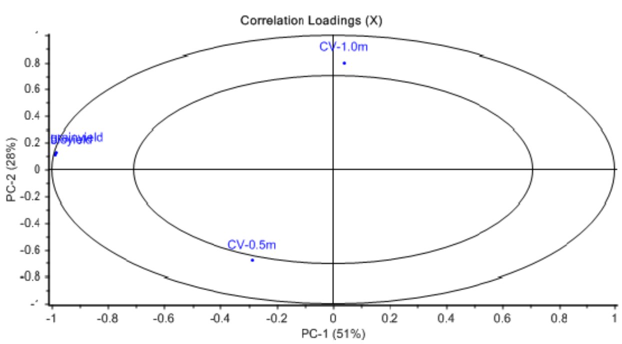


Fig. 2a and b: PCA score and loading plot analysis using EC_a and yield information with N- rate

(a)



(b)



yield productivity growth. Therefore, EC_a based tracking of N-rate and yield trends through CA can help agronomist undertake customised nutrient management program on site-specific basis.

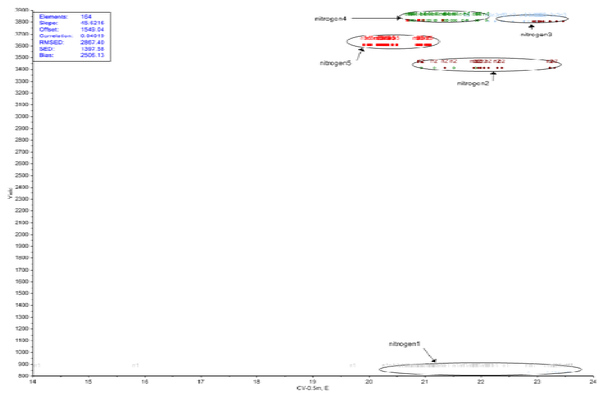
b) A similar trend having increased crop yield with the increase in nutrient rate was again observed in part2, which continued till N5 nutrient rate of application.

Similarly, crop yield increases with increase in amount of nitrogen rate in part3 and decreases after threshold of N4. Similar trend was observed in case of part1.

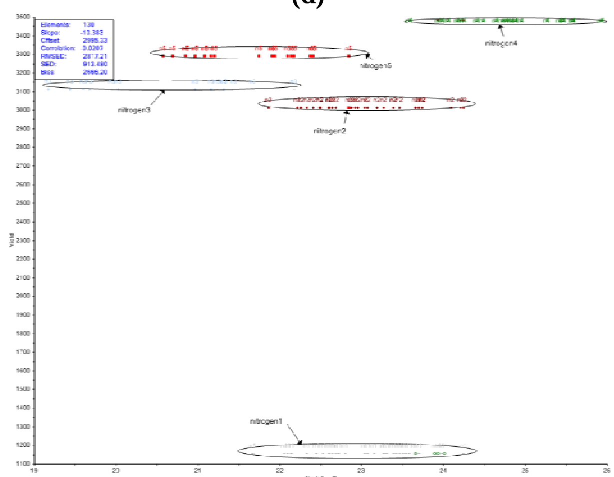
Furthermore, on sample grouping with respect to crop varieties for all plots, crop varieties show similar effect for variable nutrient rate. One of the scatter plot analysis revealed similar results when plotted between *yield* versus EC_a and observed with respect to applied *nutrient rate grouping* [(Fig. 3c & d) for part1] and [(Fig. 3e) for part2]. In these scatter plot analyzes, it was again clear that after N4 rate, there is no proportionate increase in crop yield,

Fig. 3c, d and e: Scatter plot analysis of EC_a and crop yield when observed for N-rate

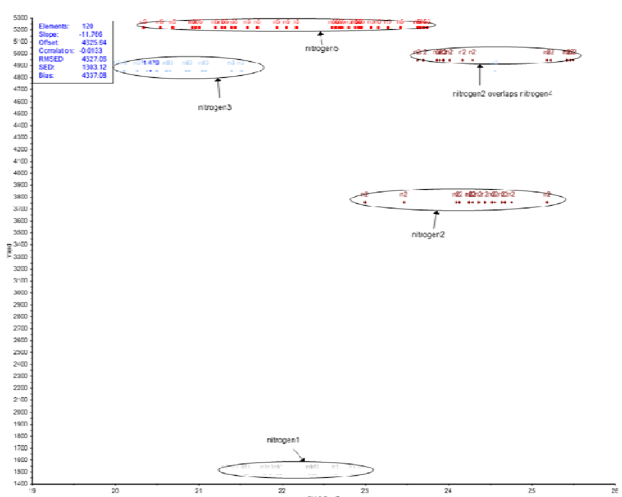
(c)



(d)



(e)



indicating over dosage done by N5 rate, which could have been avoided saving fertilizers and cost.

The scatter plot analysis indicated that for each variable nutrient rate, specific cluster of crop potential is

Fig. 4a and b: Spatial variability investigations with Hierarchical Cluster Analysis to observe matching classes between EC_a and Crop yield data

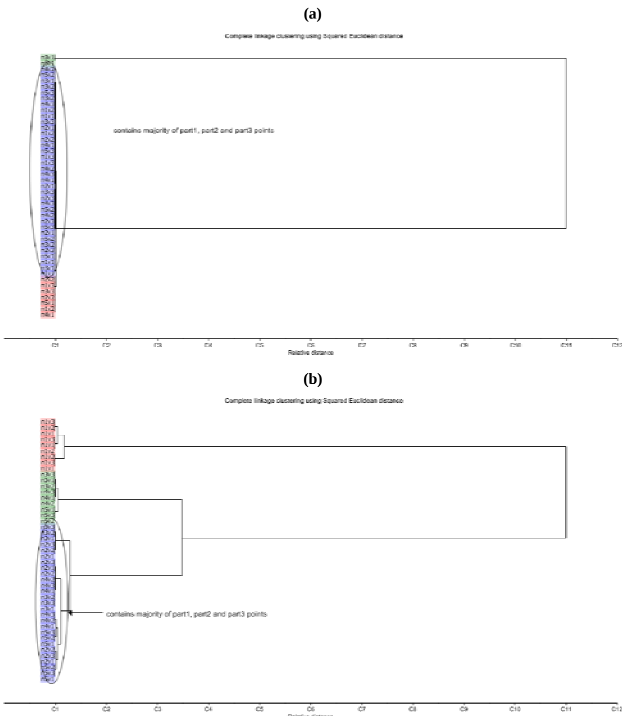
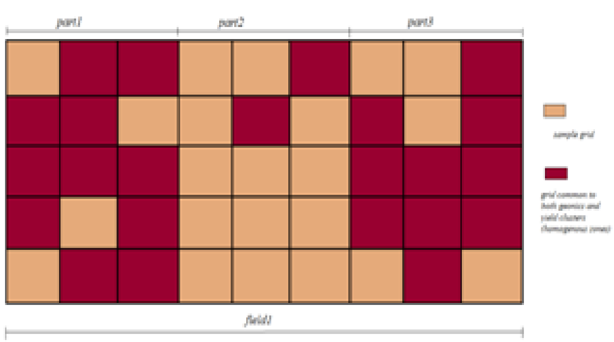


Fig. 4c: Matching patterns of EC_a data with that of yield data in field1



exhibited. These similar crop potential cluster movements can be easily tracked by the representative EC_a mapping. While systematically monitoring EC_a cluster movement, it was observed that nutrient rate management can become feasible approach using modern systems comprising of sensor measurements followed by software analysis (Fig. 3c, d & e).

Inference 3: In case of cluster analysis in field1 using EC_a and yield data, it exhibits similar number of cluster distribution for crop yield variability (Fig. 4b) and EC_a values considering measurements in the same regions (Fig. 4a).

From HCA, it is evident that the segment of part1, part2 and part3 of field1 contributes major portions of cluster classes with a clear discrimination, while identifying the homogeneity within these parts. These different cluster zones, exhibit cohesion within them i.e., having similar soil properties thus, productivity potential and hence can be given a similar soil input treatment governed by site-specific CMP of interest. A summary of analysis of grid cells depicting matching of crop yield grids to that of EC_a data clusters is shown in (Fig. 4c). However, it is visible that they are distributed in two distinct classes i.e., part 1 and part 3 as one class, whereas part 2 representing separate class.

As a result of CA, it was attempted to depict matching of patterns at a glance, observed during data analysis of entire field (Fig. 4c). When zones are delineated by CA using geonics data and yield data, the red coloured grid cells exhibit best matching patterns in agreement. Similar results were observed by Stafford *et al*. (1998) using fuzzy clustering of combine yield monitor data to divide a field into potential MZs. This shows that majority of the grids fall into the homogenous zones and show matching patterns between EC_a measurements explaining good percentage of associated yield variability, which is quite significant and encouraging. This pattern matching helps to interpret crop-yield response, which then can govern specific variable soil input rate control depending on site specific crop yield productivity potential of interest.

Thus, it was found that on-the-go; field-scale measured data of EC_a has shown encouraging interpretations associated with site-specific crop yield patterns in some parts of the field and therefore, can be considered as a surrogate measurement for field-specific crop yield response predictions.

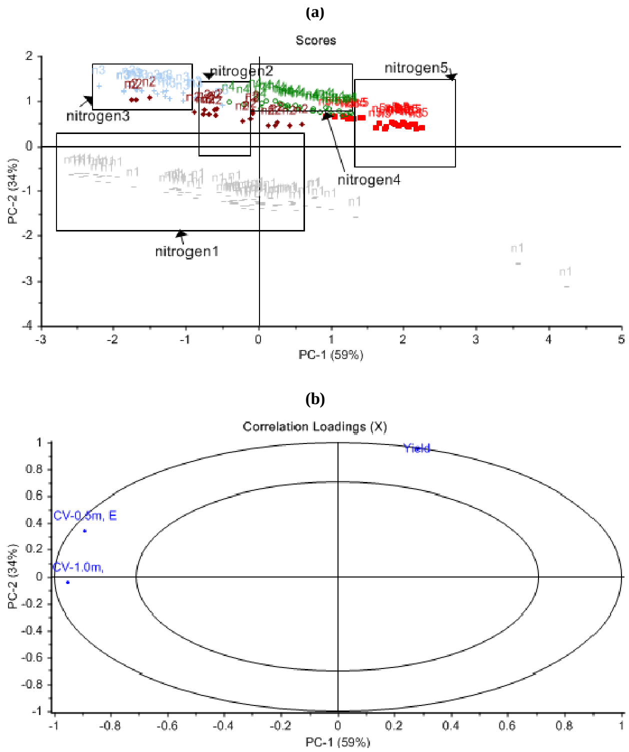
Inference 4: The field scale measured EC_a data were also studied to find variations in different zones based on the amount of nutrients applied. In total, five different quantities of nutrients namely N1, N2, N3, N4 and N5 were applied on the field1. It was attempted to determine the effect of nutrients on conductivity as well as yield response apart from classification of spatially different regions based on soil EC_a variability.

The PCA model of first variety of crop (V1), based on spatially geo-referenced EC_a measurements has shown that nitrogen amounts of N1 and N3 have the highest conductivities and are creating clusters with respect to these specific nutrient rates applied. The scores plot has clearly shown the creation of clusters i.e., homogenous zones with respect to different nitrogen rates applied. This has been confirmed by the conductivity based scores plot (Fig. 5a & b).

The PCA model, for second variety of crop (V2), based on EC_a, has shown that nitrogen amount N2 has the highest conductivity.

Similarly, PCA model for third variety of crop (V3) had shown that nitrogen amount N4 has the highest conductivity; the same is confirmed by plot.

Fig. 5a and b: PCA analysis of EC_a, nutrient rate and yield for crop variety V1



While analysing PCA model for all three varieties, it was found that N5 has shown agreement with highest yield trend and it increases towards right hand side of loading plot in the vertical direction. This analyzes helped to understand that EC_a based measurements are able to distinguish five different clusters that were applied with five different N-rates, resulting in five variations in crop yield. Besides above observations, scores plot has shown cluster formations in remaining cases of variable n-rates application over entire field.

Referring graphical representation of PCA analysis, for all three varieties of paddy, based on spatial measurement of soil EC_a data for two soil depths i.e., 0-75 cm and 0-150 cm, it can be observed that each variable rate has separate cluster as shown in scores plot. It suggests that variable rate of nutrients can be classified, managed and monitored using bulk conductivity monitoring. While observing score plots for all three varieties, it was found that different clusters of conductivity belong to specific crop yield response, that means yield cluster have matching EC_a response characteristics. Hence, variable soil input rate can be customised and monitored in conjunction with EC_a measurements and cluster analysis. Further, it was analysed that increase in input rate such as N4 and N5 has indicated distinct cluster patterns indicating highest crop yield irrespective of crop variety.

Inference 5: The geonics measured EC_a data were also used to differentiate amongst three fields in order to observe any natural clusters so as to treat and govern management zone with variable soil input rate control between fields. Following are the inferences drawn, when all fields were analysed altogether with PCA at once.

a) Cluster formations have been observed for each field in the PCA scores plot. The observed clusters can be seen (Fig. 6a) that are differentiated by colouring scheme that corresponds to different fields as follows:

Red: field1, Green: field2, Blue: field3.

Sample grouping was carried out with respect to field EC_a ranges; Principal Component (PC) 1 of measured data set explains 100% variance of the soil data in the selected fields. This knowledge of cluster formation was used to represent customized zones, which can be treated with specific soil input rates to produce bio-eco farming returns. It was also seen through lab analyzed data that field 1 and field 2, which are close to each other in scores plot are having nearly similar soil texture characteristics compared to field 3, which is clearly discriminated as a separate cluster.

b) In Fig. 6a, it has been observed that the properties of field 1 and field 2 are similar to some extent as compared to that of field 3 cluster.

c) The study of loadings plot (Fig. 6b) reveals that conductivities measured at two depths lie on the outer ellipse and hence is responsible to explain maximum variance of the field. But, they do not exhibit a strong positive correlation and are at two extreme ends with each other meaning that they represent two different sub-soil spatial variability at two different depths covering different sub-soil volumes, hence it becomes necessary to measure the conductivities at both the depths to investigate soil depth characterization at two depths covering different volumes and sub-soil characteristics. The characterization of specific root-zone depth is of interest in case of crop-yield response studies, to understand crop specific rooting zone properties on a site-specific manner.

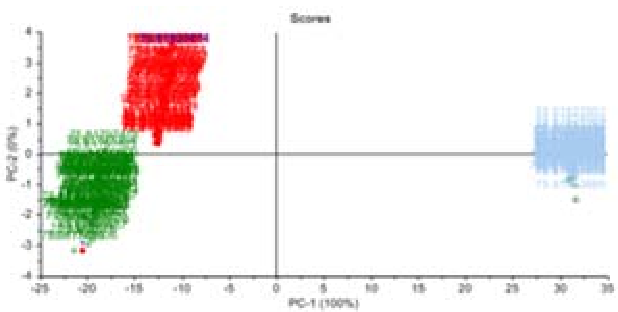
d) A comparative study of scores plot and the loadings plot (Fig. 6a & b) explains that the significant soil attribute for describing the characteristics of field 1 and field 2 is apparent conductivity mapped for (0-150 cm) depth, while that of field 3 is conductivity for (0-75 cm) depth. This suggests that for field 3, the soil spatial variability is better explained and sub-soil spatial characteristic is positively correlated with EC_a measurements for (0-75 cm) depth, whereas in case of fields 1 and 2, the conductivity measurements for (0-150 cm) is more meaningful to govern further investigations.

It was also observed (Fig. 7) with the help of correlation loadings plot that conductivity measured for (0-75) cm is significant in explaining the variabilty of crop yield as compared to EC_a measured for (0-150) cm.

Analysis of correlation and different clustering techniques for delineation of MZs: The correlation between soil EC_a and physico-chemical properties of the soil was evaluated based on spatial variability surface maps generated using kriging and co-kriging interpolation (Fig. 8).

Fig. 6a and b: PCA analysis for combined three fields to observe classification of zones based on EC_a

(a)



(b)

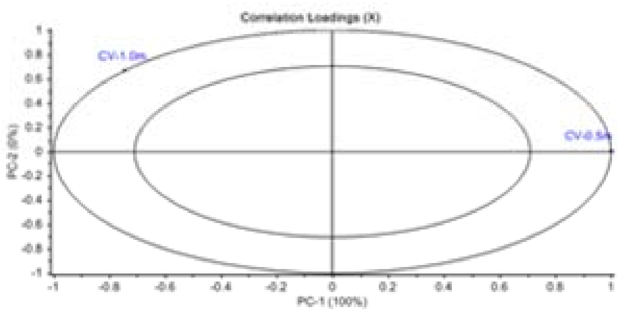
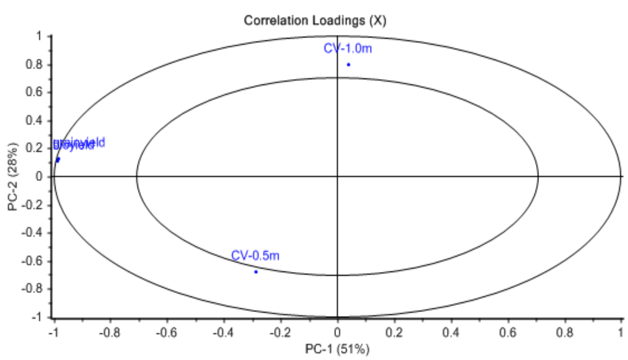


Fig. 7: Correlation loading plot to study EC_a measured at different depths with Crop yield

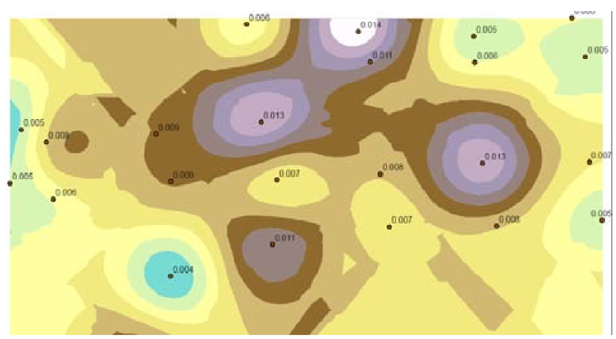


Strong correlations were identified after analyzing the relation between measured EC and soil attributes (Table II).

These interactions and intricacies amongst the various soil attributes are limiting factors for the use of simple correlation analysis in the interpretation of these data with respect to crop yield. Therefore, PCA, a dimensionality reduction technique without loss of relevant information, was then used to reduce the number of variables and to generate MZs. By applying PCA, scores were obtained that are projection of actual data on PC vector space. Variance values of each PC are shown in Table III. Significant PCs were selected among the total set of original variables which together explained 93.93% of the total cumulative variance of those data.

Fig. 8: Geo-spatial variability of (a) soil Electrical conductivity (b) other physiochemical properties

(a)



(b)

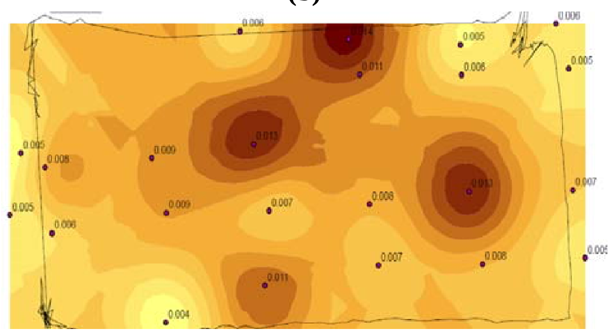
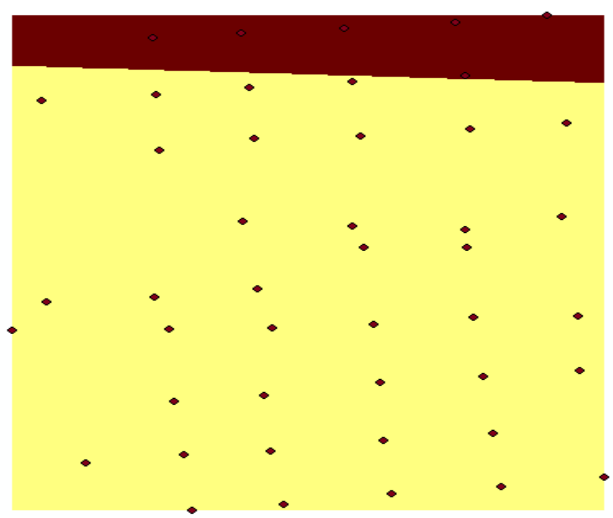


Fig. 9: EC_a delineated zones



Both the data groups were separately clustered with fuzzy-c means algorithm. The integrated script of PCA and Fuzzy c-means written in MATLAB program produces the results (Figs. 9 & 10), where it shows delineated zones of EC_a and yield having matching pattern.

In order to delineate precise MZs, optimal number of clusters using fuzzy clustering process was obtained and verified using fuzzy performance index (FPI) and normalized classification entropy (NCE). The process was carried out for both conductivity and yield data (Figs. 11 & 12, respectively).

Table I: Descriptive Statistics of geo-sensed, field scale measured EC_a

Statistics	Soil Information	Mean	Max	Min	Range	SD	Variance	RMS	Median
	CV-1.0m (Ec_a for 0-150cm)	23.45255	27.583	20.943	6.639999	1.841426	3.390848	23.52313	23.169
	CV-0.5m (Ec_a for 0-75cm)	31.22414	231.44	19.248	212.192	43.4697	1889.615	53.12786	21.904
	Grainyield	1431.17	2282.5	345.344	1937.156	586.1918	343620.8	1544.096	1551.348
FIELD1	Bioyield	3318.218	5217.93	836	4381.93	1273.822	1622622	3549.243	3483.11
	CV-1.0m (Ec_a for 0-150cm)	18.12677	23.894	14.128	9.765999	2.097522	4.3996	18.24748	17.996
FIELD2	CV-0.5m (Ec_a for 0-75cm)	12.27878	16.733	7.436	9.297	1.928088	3.717523	12.42916	12.358
	CV-1.0m (Ec_a for 0-150cm)	15.88289	19.786	11.193	8.592999	1.654272	2.736615	15.9687	16.193
FIELD3	CV-0.5m (Ec_a for 0-75cm)	62.60425	66.713	56.517	10.196	1.543256	2.381638	62.62354	63.002

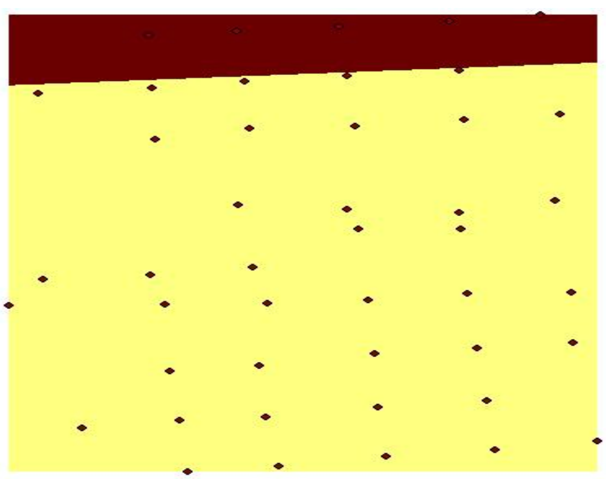
Table II: Fitness Measures Statistics of EC and physiochemical soil properties and yield

EC-Relation	R-square	RMSEC	RMSEP	SEC	SEP	Correlation
Moisture	0.937595	0.000828	0.000856	0.000834	0.000862	0.968295
Temperature	0.034019	0.003257	0.003415	0.003280	0.003439	0.184443
Real Dielectric	0.947069	0.000763	0.000780	0.000768	0.000786	0.973175
pH	0.002080	0.003311	0.003400	0.003334	0.003424	0.045603
Organic Matter	0.566271	0.006381	0.006427	0.006397	0.006429	0.75251
Clay	0.479334	0.008472	0.008618	0.008571	0.008614	0.69234
Yield	0.506118	0.006287	0.006318	0.006275	0.006348	0.71142

Table III: Principal component vs. Cumulative Variance

Principle Component	Variance Cumulative
1	0.6612
1-2	0.9393
1-3	1.0000
1-4	1.0000
1-5	1.0000

Fig. 10: Yield delineated zones



The optimal number of clusters for each computed index was observed when the index is at the minimum, representing the least membership sharing (FPI) or greatest amount of organization (NCE) as a result of the clustering process. The minimum FPI and NCE were obtained with two clusters for the present study area, which is also coincident with other analysis in this area discussed earlier including HCA. The successful validation by FPI and CE indices implies least membership sharing and greatest amount of organization therefore, demonstrating distinct classes or MZs.

Fig. 11: FPI and CE indices of conductivity

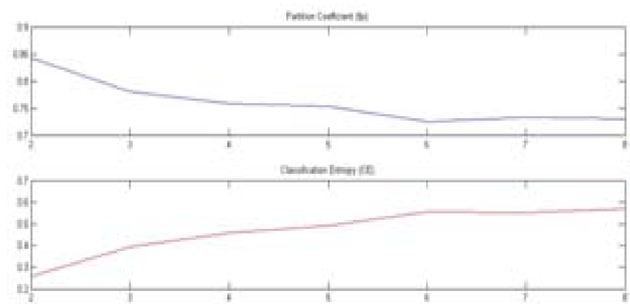
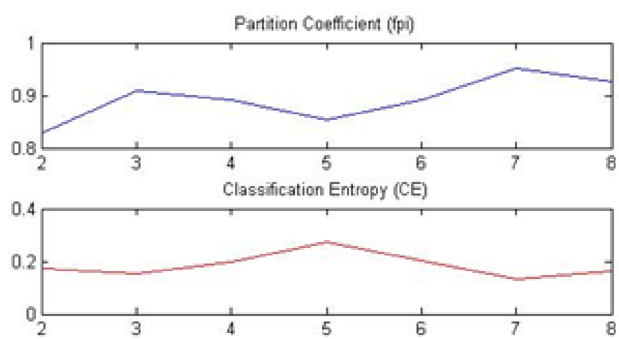


Fig. 12: FPI and CE indices of Yield



Many researchers use yield maps for generating productivity MZs. However, it has been debated time and again that it involves considerable time and expense and the cropping inputs necessary to optimize productivity and minimize environmental impacts which can be derived only if factors contributing to the observed spatial crop patterns is known. Yield maps alone do not provide this information nor do they by themselves provide the information necessary to differentiate edaphic, anthropogenic, biological and meteorological factors influencing crop yield and spatial crop patterns. Also yield monitors have not been developed for all crop varieties and it needs appropriate calibration compensations in case of real-time monitoring, which are highly subjective. In contrast, EC_a is an easy-to obtain, rapid and low-cost method of soil and landscape measurements to provide productivity MZs. In order to test whether customized management zones (CMZs) from easily obtained EC_a would be similar to those created from the more difficult, time-consuming and expensive yield-mapped data, clusters identifying spatial patterns of yield and conductivity were plotted along with respective location co-ordinates (Figs. 13 & 14).

A visual inspection of these maps indicated a good agreement between both. This relationship between EC_a and yield was previously reported by Lund *et al*. (2000).

It can be observed from overlapped results (Fig. 15) that clusters of yield and conductivity are matching, which indicated that yield interpretations can be made from easily obtained EC_a measurements and thus, can be used to create CMZs to govern variable soil input rates. This verifies the utility of EC_a in identifying the areas within field that correspond to homogeneous soil fertility zones. It can also be implied that directed soil sampling can be undertaken, depending upon desired farming guidance and/or agronomic resource application. Based on the studies for another 2-3 cycles from the same field, development of within field micro-management zones can be established in order to provide optimal amount of targeted soil input rates. This type of customized rational soil input treatment approach results in ensuring safe food chain supply system, which is free from excess fertilization imbalances and can generate stable inroads to country's economic development owing to fertilizer being core industry influencing GDP seriously, apart from protecting natural infrastructure for the survival of mankind as well as micro-organisms.

Further, to visualize homogeneity among the clusters, they were imported to ArcCatalog and later to Arc map. Kriging was applied on these clusters and exponential semivariogram model (Lags=7, Neighbours=5) was built and thus, delineated MZs were designed for each field (Fig. 16). The point with the same color (i.e., cluster) lies in the single MZs. Only the overlapping clusters are distributed in the various MZs as they share the properties of both clusters. Krigged and co-krigged based interpolated spatial maps were generated after selecting appropriate neighbourhood criteria, cross validation, variogram model, using different arc map and geostatistical options of ArcInfo program (ESRI ArcGIS software).

While creating geospatial maps, prediction equations were used to generate surfaces for unknown values using known ones, and standard error estimations were also computed. In this process, various parameters such as nugget, sill, lag size, standard errors were noted (Table IV).

Further, scatter plots and groupings of variables are becoming increasingly common for PF service providers to observe similarity and regression correlation within paired data. When this is applied to geo-referenced EC_a and yield data sets and plotted with respect to latitude and longitude, resulted producing typical matching patterns (Fig. 17).

Fig. 13: Yield clusters along with latitude and longitude

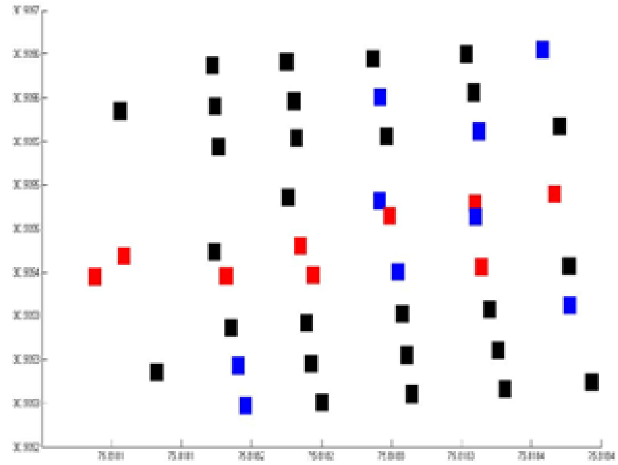


Fig. 14: Conductivity clusters along with latitude and longitude

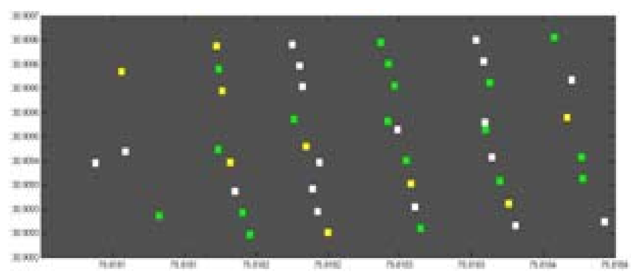
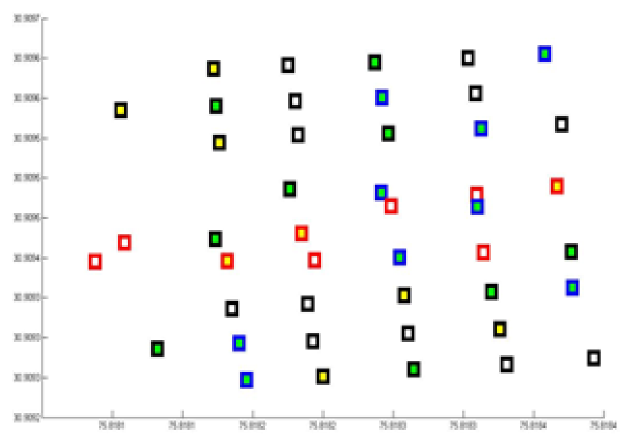


Fig. 15: Yield and conductivity overlapped clusters along with latitude and longitude



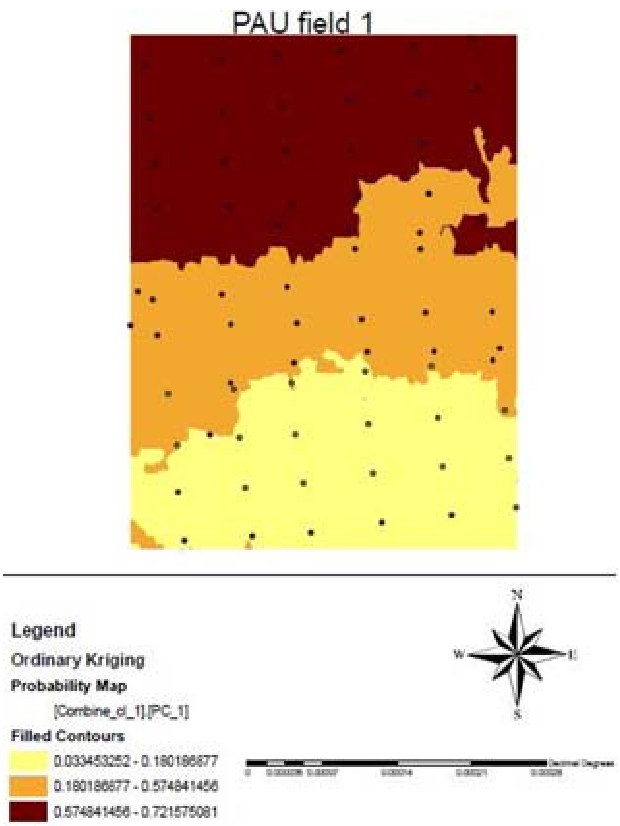
CONCLUSION

The present paper demonstrated how modern techniques such as PCA, Fuzzy c-means and HCA could effectively be used to group correlated and similar soil properties representative i.e., EC_a into unique groups, for applying variable soil input rates using site-specific footprint

Table IV: Semivariogram and covariance modeling parameters for soil EC_a and other physiochemical attributes

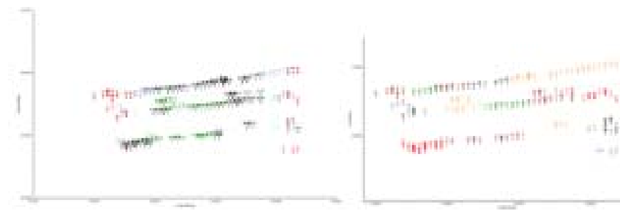
Soil attribute	Nugg -et effect	Partial Sill	Lag size	Standardized errors				Model
				Root-Mean-Square	Mean	Average Standard Error	Root-Mean-Square Standard-ized	
EC	0.15	0.25427	0.000014143	0.00271	0.01145	0.002941	0.9269	exponential
Moisture	0.00122	0.31393	0.007814143	0.05095	0.02577	0.05223	0.9832	exponential
pH	0.12045	0.28576	0.000014267	0.5535	-0.06501	0.5868	0.9436	Gaussian
OM	3.52	0.17689	0.005314143	0.3263	0.04456	0.7789	0.8763	Linear
Clay	8.62	0.87531	0.000914143	0.0448	0.3367	0.3421	0.5432	Linear

Fig. 16: Delineated Management zones for Fields 1, 2 and 3



which uses customized zone based management principles. This helps to interpret/predict crop yield response owing to spatial variations, and apply improved as well as optimal management practices for sustainability and bio-eco returns. The application of PCA and other clustering techniques in the present studies have identified matching patterns between geonics measured EC_a data and associated crop yield variations. Moreover, EC_a information of soil was analysed for soil variability with respect to variable nutrient rates applied in the experimental field of PAU, Ludhiana, India. It can be concluded that the soil spatial variations measured in terms of EC_a can be termed as a surrogate indicator of sub-soil spatial variability that affects crop yield variations and hence can be employed in conjunction with cluster and PCA techniques to govern customized soil input rate management practices. In present studies, variable nutrient rates applied over entire field1 was also correlated

Fig. 17: Typical plot of soil EC and yield in agreement



with measured EC_a through clustering technique and PCA score plots for a specific location and depth. In one of the studies, EC_a, crop bio/grain yield data were analysed with regression technique and results have indicated that specific field characteristics are required to be taken into consideration, while developing site-specific calibration equation for a parameter of interest. It was also observed in the reported procedure that there is no limitation on number of properties to be considered, while delineating zones, as both, PCA and clustering techniques work on multidimensional data, which otherwise is the limiting factor in other spatial variability analysis methods such as krigging and co-krigging etc. It was seen that different CA techniques have produced similar results while delineating MZs. Such chemometric techniques in the field of PF have a long way to go in order to ensure sustainable agriculture, improving quality of life of farmers, and ever continuing growth.

Acknowledgment: Authors express their deep sense of gratitude towards Dr. Pawan Kapur, Director, CSIR-CSIO, Chandigarh, for the guidance and encouragement provided from time to time in undertaking this development. The authors also feel highly obliged towards Dr. Manjeet Singh, research scientist working in the farm and machinery department of PAU, Ludhiana, (INDIA), for providing research farm fields and assistance in conducting lab analysis of soil samples, which helped the authors to accomplish this work.

REFERENCES

Fridgen, J.J., N.R. Kitchen and K.A. Sudduth, 2000a. Variability of soil and landscape attributes within sub-field management zones. *Proc. Int. Confr. Prec. Agric.*, Vol. 5. Bloomington, 2000

Fridgen J.J., C.W. Fraisse, N.R. Kitchen and K.A. Sudduth, 2000b. Delineation and Analysis of site-specific management zones. *2nd Int. Confr. Geospatial Info. Agric. Forestry.* Lake Buena Vista, Florida 10-12 Jan. 2000

Heiniger, R.W., R.G. McBride and D.E. Clay, 2003. Using Soil Electrical Conductivity to Improve Nutrient Management. *Agron. J.*, 95: 508–519

Jaynes, D.B., J.M. Novak, T.B. Moorman and C.A. Cambardella, 1994. Estimating Herbicide Partition Coefficients from Electromagnetic Induction Measurements. *J. Envirn. Qual.*, 24: 36–41

Jaynes, D.B., T.S. Colvin and J. Ambuel, 1995. Yield mapping by electromagnetic induction. *2nd Int. Confr. Site-Specific Management for Agric. Systems*, pp: 383–394. ASA-CSSA-SSSA, Madison, Wisconsin, USA

Jose, P.M. and C.N. de Castro, 2008. Establishing Management Zones using Soil Electrical Conductivity and other Soil Properties by the Fuzzy Clustering Technique. *Sci. Agric.*, 65: 567–573

Kitchen, N.R., K.A. Sudduth, D.B. Myers, S.T. Drummond, and S.Y. Hong, 2005. Delineating productivity zones on claypan soil fields using apparent soil electrical conductivity. *Comp. Electr. Agric.*, 46: 285–308

Luchiari, Jr., A., J. Shanahan, D. Francis, M. Schlemmer, J. Schepers, M. Liebig, A. Schepers and S. Payton, 2000. Strategies for establishing management zones for site specific nutrient management. *Inter. Confr. Precision Agric.*, Vol. 5. Bloomington, 2000

Lund, E.D., C.D. Christy and P.E. Drummond, 2000. Using yield and soil electrical conductivity (EC) maps to derive crop production performance information. *In*: Robert, P.C., R.H. Rust and W.E. Larson (eds.), 5th *Inter. Confr. Precision Agric.* Minneapolis, MN, 16–19 July 2000. ASA, CSSA, and SSSA, Madison, WI

Lund, E.D., M.C. Wolcott and P. Hanson Glenn, 2001. *Applying Nitrogen Site-specifically using Soil Electrical Conductivity Maps and Precision Agriculture Technology*. 2nd International Conference of Science and Policy, Potomac MD, October 14-18, 2001

McBride, R.A., A.M. Gordon and S.C. Shrive, 1990. Estimating forest soil quality from terrain measurements of apparent electrical conductivity. *Soil Sci. Soc. America J.,* 54: 290–293

McNeill, J.D., 1992. *Rapid Accurate Mapping of Soil Salinity by Electromagnetic Ground Conductivity Meters*, pp: 209–229. Advances in Measurement of Soil Physical Properties: Bringing Theory into Practice', Special Publication 30. SSSA, Madison, WI, 1992

Moral, F.J., J.M. Terron and F.J. Rebollo, 2011. Site-specific management zones based on the Rasch model and geostatistical techniques. *Computers Electronics Agric.*, 75: 223–230

Ouyang, D., A.F. MacKenzie and M. Fan, 1998. Phytotoxicity of banded urea amended with triple superphosphate and potassium chloride. *Agron. J.*, 90: 734–739

Rhoades, J.D., D.L. Corwin and S.M. Lesch, 1999. Geospatial measurements of soil electrical conductivity to assess soil salinity and diffuse salt loading from irrigation. *In*: Corwin, D.L. (ed.), *Assessment of Non-Point Source Pollution in the Vadose Zone*, pp: 197–215. Geophysical Monograph 108, American Geophysical Union, Washington, DC, 1999

Stafford, J.V., R.M. Lark and H.C. Bolam, 1998. Using yield maps to regionalize fields into potential management units. *In*: Robert, P.C. (ed.), 4th *Inter. Confr. Precision Agric*., pp: 225–237. ASA, CSSA, SSSA, Madison, Wisconsin, 1998

Vrindts, E., M.D. Reyniers, P. Marc, B.H. Frankinet, M.-F. Destain and J. De Baerdemaeker, 2003. *Analysis of Spatial Soil, Crop and Yield Data in a Winter Wheat Field*. 2003 ASAE Annual International Meeting, ASAE., Nevada, USA

Williams, B. and D. Hoey, 1987. The use of electromagnetic induction to detect the spatial variability of the salt and clay content of soils. *Australian J. Soil Res.*, 25: 21–27

(Received 03 February 2012; Accepted 06 March 2012)