THE INTERNATIONAL SOCIETY OF PRECISION AGRICULTURE PRESENTS THE
14th INTERNATIONAL CONFERENCE ON
PRECISION AGRICULTURE
JUNE 24th-27th 2018
Le Centre Sheraton • Montreal, Quebec • Canada

Data Clustering Tools for Understanding Spatial Heterogeneity in Crop Production by Integrating Proximal Soil Sensing and Remote Sensing Data

Md Saifuzzaman[1], Viacheslav Adamchuk[1], Hsin-Hui Huang[1], Wenjun Ji[2], Nicole Rabe[3], Asim Biswas[4]

[1]Dept. of Bioresource Engineering, McGill University, Ste-Anne-de-Bellevue, Quebec, Canada

[2]Dept. of Soil and Environment, Swedish University of Agricultural Sciences, Skara, Sweden

[3]Environmental Management Branch, Ontario Ministry of Agriculture, Food and Rural Affairs, Guelph, Ontario, Canada

[4]School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada

A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada

Abstract.

Remote sensing (RS) and proximal soil sensing (PSS) technologies offer an advanced array of methods for obtaining soil property information and determining soil variability for precision agriculture. A large amount of data collected using these sensors may provide essential information for precision or site-specific management in a production field. In this paper, we introduced a new clustering technique was introduced and compared with existing clustering tools for determining relatively homogeneous parts of agricultural fields. A DUALEM-21S sensor was used, along with high-accuracy topography data, to characterize soil variability from three agricultural fields in Ontario, Canada. Sentinel-2 data were used for measuring bare soil and historical vegetation indices (VIs). The custom Neighborhood Search Analyst (NSA) data clustering tool was implemented using Python. In this NSA algorithm, part of the variance of each data layer is accounted for by subdividing the field into smaller relatively homogeneous areas. The algorithm was illustrated using field elevation, shallow and deep EC$_a$, soil pH, and several VIs.

Keywords.

Proximal soil sensing, remote sensing, spatial data, clustering techniques, management zones.

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 14th International Conference on Precision Agriculture. EXAMPLE: Lastname, A. B. & Coauthor, C. D. (2018). Title of paper. In Proceedings of the 14th International Conference on Precision Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture.

Introduction

Management zone delineation using remote sensing (RS) and proximal soil sensing (PSS) data is becoming important for the assessment of soil property and characterizing variability in precision agriculture (Shatar & McBratney, 2001; Fridgen et al., 2004; Dhawale et al., 2014; Albornoz et al., 2018). In the delineation process, high-resolution data from these sensing technologies, together with quantitative methods, is used to infer the spatial pattern of soil heterogeneity (Deng et al., 2003; Adamchuk et al., 2004; Cohen et al., 2013; De Benedetto et al., 2013). To obtain information on the spatial pattern of soil and to produce the thematic soil maps of a field for understanding agronomic and yield-limiting factors, high density and multivariate data analysis were used to determine a solution by isolating homogeneous field areas and potential management zones (Vrindts et al., 2005; LI et al., 2007; Cressie & Kang, 2010; Adamchuk et al., 2011, Dhawale et al., 2016).

Multivariate data clustering techniques are imperative to achieve significant benefits from identifying and understanding soil variability within a production field (Burrough et al., 1997; Ruß & Brenning, 2010). Among the multivariate data analysis techniques, clustering techniques are most commonly used. Various indices in the non-hierarchical cluster analysis from fuzzy c-means (FCM) and from K-means are among the common clustering methods used for data mining (Gui-Fen et al., 2007; Panda et al., 2012). Due to the fuzziness of C-means and K-means, and several other limitations (i.e., create boundary pixels and each cluster object belongs in one or more groups) in the isolation process (Albornoz et al., 2018), this study attempts to provide a multivariate clustering tool to represent unique thematic maps and zonal boundaries based on the homogeneity of the agricultural field. Most of the clustering algorithms used in zone delineation do not handle high-density data files with multiple variables (Viscarra Rossel et al., 2011; Córdoba et al., 2016). The agricultural scientist and farmers often face various challenges for variable rate operations due to fragmented management zones, which are commonly produced by a clustering technique (Albornoz et al., 2018). The objective of this study was to present the process for developing a new and enhanced clustering technique to better understand soil variability in an agricultural field and compare them with commonly used ones.

Material and methods

Experimental sites and data description

Three agricultural fields of varied sizes from Woodrill Farms in Ontario, Canada were mapped using both RS and PSS sensors (Table 1 & Fig 1). Elevation data was collected by Real-Time Kinematic (RTK) Global Navigation Satellite Systems (GNSS) from the agricultural fields (Table 2). Slope, aspect (sin), and topographic wetness index (TWI) variables were derived from a digital elevation model (DEM) of the study sites. Dualem 21s was used to collect apparent electrical conductivity (EC_a) of four different depths: HCP1 – 0-1.6 m, PRP1 – 0-0.5 m, HCP2 – 0-3 m, and PRP2 – 0-1 m (Table 3). Potential outliers and null values of the PSS measurements were removed in the preprocessing steps, and about 15% of the PSS data was removed. Ordinary Kriging interpolated maps were generated from the PSS measurements in ESRI ArcGIS software. Various geospatial (e.g., rectification, point data extraction etc.) and digital remote sensing data processing (e.g., radiometric correction, stitching, stack bands etc.) steps were followed, and these enhanced the data quality for further analysis. Finally, the text data file was generated to store the sensor variables and sensor-derived variables for use in the clustering process.

Table 1. Three agricultural fields in Ontario, Canada.

Field ID	Area (ha)	Target crops
WH	39.55	Soybean/Wheat
LD	21.00	Soybean
RB	75.00	Soybean/Wheat

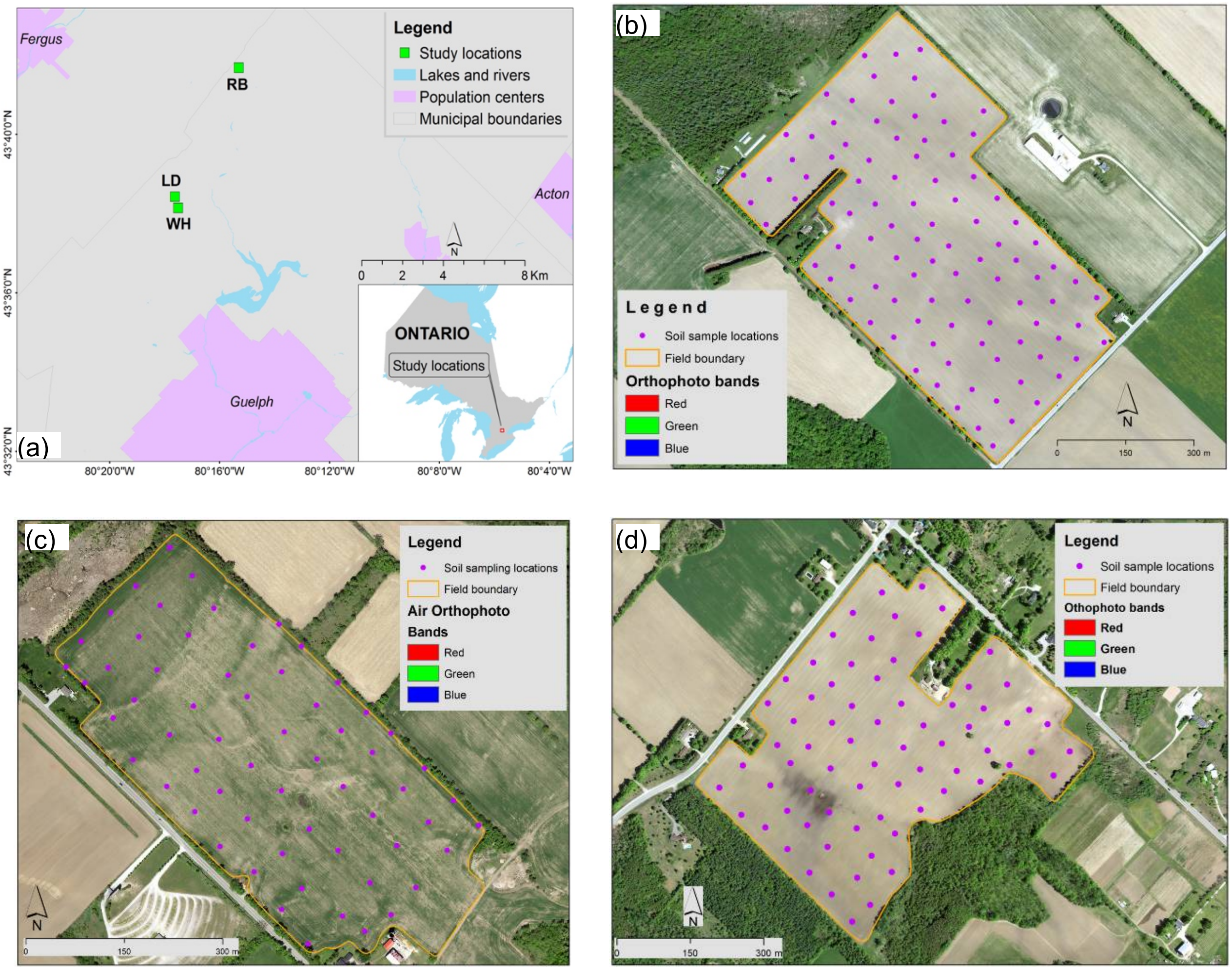

Fig 1. (a) Three Woodrill farms in Ontario: WH field boundary with soil sample locations (b), LD field boundary with soil sample locations (c), and RB field boundary with soil sample locations (d).

Table 2. Summary statistics of elevation data from RTK sensor in the three agricultural fields

Field ID	# of measurements	Elevation (m)					
		Min	Median	Max	Range	STD	Mean
WH	28493	372.06	378.07	384.54	12.48	2.33	378.21
LD	7110	332.70	344.86	354.17	21.47	5.76	343.95
RB	20813	358.41	367.67	372.16	13.75	3.63	366.64

Table 3. Summary statistics of DUALEM 21s sensor variables from the three agricultural fields

Field ID	# of measurements	Sensor configuration	Apparent soil electrical conductivity (EC$_a$), mS/m					
			Min	Median	Max	Range	STD	Mean
WH	20129	HCP1	4.00	12.28	25.28	21.28	1.69	12.51
LD	6931		2.58	6.90	16.08	13.50	1.55	6.96
RB	18524		1.70	9.00	17.98	16.28	2.81	9.13
WH	20129	PRP1	4.68	7.92	22.24	17.56	1.60	8.15
LD	6931		0.72	4.44	14.12	13.40	1.38	4.55
RB	18524		0.00	3.53	16.80	16.80	2.86	4.40
WH	20129	HCP2	7.42	10.46	24.42	17.00	1.79	10.83
LD	6931		0.50	4.44	14.44	13.94	1.85	4.61
RB	18524		2.50	8.45	14.99	12.49	2.65	8.22
WH	20129	PRP2	5.42	9.10	23.92	18.50	1.75	9.37
LD	6931		1.08	4.68	14.60	13.52	1.50	4.75
RB	18524		0.14	5.10	15.00	14.86	2.96	5.64

High spatial and spectral resolution images were used for analyzing bare soil and historical vegetation characteristics (Table 4). Among the vegetation indices (VIs), NDVI maps from Sentinel-2 data were found to be more suitable and were used for the clustering process (Roberts et al., 2011; Viña et al., 2011). Orthophoto and Sentinel-2 four red-edge bands were used for visual interpretation with zonal thematic maps.

Table 4. Remote sensing data characteristics and its sources

Satellite sensor	Spectral bands	Pixel (m)	Central Wavelength(nm)	Imaging date	Source
OrthoPhoto	B, G, R, NIR	0.2	-	May 23, 2015	OMAFRA/OMNRF*
Sentinel-2	2(B), 3(G), 4(R), 8(NIR)	10.0	494, 560, 665, 834	July 21, 2017	Planet Labs
	5,6,7 (Red-edge 1,2 &3)	20.0	704, 740, 781	July 21, 2017	Planet Labs

Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) & Ontario Ministry of Natural Resources and Forestry (OMNRF)

Data clustering algorithms

Fuzzy C-means in the management zone analyst (MZA) (USDA, 2000) were used for generating the normalized classification entropy (NCE) and fuzziness performance index (FPI) of the maximum five zones. The *K*-means algorithm in Python data library also was used to generate 5, 15, and 25 clusters, and to find cluster centers using the sum of square distances of all data points and the number of cases in each cluster.

The proposed data clustering method, called neighborhood search analyst (NSA), resulted in the algorithms shown in Fig 2. The processing steps and formula are adopted from NSA written in MatLab code (Dhawale et al., 2014). To construct an objective function to be optimized through the data grouping process, the mean squared error (MSE) was calculated for each individual data layer k according to:

$$MSE_k = \frac{\sum_{j=1}^{m}\sum_{i=1}^{n_j}\left(X_{ij} - \overline{X}_j\right)^2}{N - m} \qquad (1)$$

where, X_{ij} is a sensor value for the i_{th} grid cells within the j_{th} group;
$\overline{X}_j$ is the mean of j_{th} group;
N is the total number of grid cells;
m is the number of groups;
n_j is the number of grid cells within the j_{th} group.

It should be noted that the difference between the total number of grid cells and the number of groups can be determined:

$$N - m = \sum_{j=1}^{m} \left(n_j - 1 \right) \tag{2}$$

Since the algorithm initially assumes that all data elements belong to the same group number 1 was, named "the rest of the field". *$MSE_k(m=1)$* represents the variance of k_{th} data layer across the entire field. Considering that the area of the field is substantially greater than the area of a grid cell, *$MSE_k(m=1)$* can be called Farthest Distance Variance (*FDV_k*). In such a situation, the portion of data variance accounted for by distributing *N* grid cells among *m* groups can be calculated according to:

$$R_k^2 = 1 - \frac{MSE_k}{FDV_k} \tag{3}$$

where *$MSE_k(m=1)$* can be called Farthest Distance Variance (*FDV_k*).

The maximum value of R^2_k can be obtained when MSE_k is as small as possible and it is approaching 1 when the number of groups increases. Since the result can be considered less favorable if at least one data layer *k* is not adequately accounted for, it is reasonable to employ the integration operator OR instead of the more common AND. This excludes the need to assign a weight factor to each individual data layer when adding corresponding MSE_k estimates. In mathematical terms, this would mean that the product of all R^2_k should be maximized. Therefore, the objective function (*OF*) was defined as:

$$OF = \prod_{k=1}^{K} R_k^2 \tag{4}$$

where *K* is the number of PSS data layers.

In this research, the smallest number of data elements that could be grouped was assigned to be a nine (3 x 3) grid cell square window. Therefore, the maximum accountable variance is the variance of PSS measurements between immediate neighbors. The Shortest Distance Variances (*SDV_k*) can be found using:

$$SDV_k = \frac{1}{w} \sum_{j=1}^{w} \sum_{i=1}^{9} \frac{\left(X_{ij} - \overline{X}_j \right)^2}{8} \tag{5}$$

where *w* is the total number of 3x3 square windows of grid cells.

Since SDV_k represents the smallest MSE_k value, the maximum value of R^2_k is calculated as:

$$R_{k\,\max}^2 = 1 - \frac{SDV_k}{FDV_k} \tag{6}$$

This $R^2_{k\ max}$ parameter can range between 0 and 1. It is equal to 0 when data layer *k* is either uniform, or highly variable so that $SDV_k = FDV_k$. In such a case, the data layer should not be able to affect changes in the OF. Alternatively, when $R^2_{k\ max}$ is close to 1, the data layer has a strong spatial structure ($SDV_k << FDV_k$) and OF must be sensitive to the change of MSE_k corresponding to that particular data layer.

In mathematical terms, this goal can be achieved by multiplying all R^2_k values raised to $R^2_{k\ max}$ power:

$$OF = \prod_{k=1}^{K} R_k^{2\ R_{k\max}^2} = \prod_{k=1}^{K} \left(1 - \frac{MSE_k}{FDV_k} \right)^{\left(1 - \frac{SDV_k}{FDV_k} \right)} \tag{7}$$

The resultant OF indicates the overall quality of grid cell groupings. It varies from 0 to 1 and approaches high values when every spatially structured layer of PSS measurements is separated among spatially continuous groups of grid cells with minimum internal group variance. Such groups represent different combinations of average PSS measurements obtained with different sensors that diverge from average field conditions. To facilitate the formation of grid cell groups that would maximize the OF, the NSA algorithm was implemented in this study using Python v3.6 (created by Guido van Rossum and managed by Python Software Foundation, Delaware, USA).

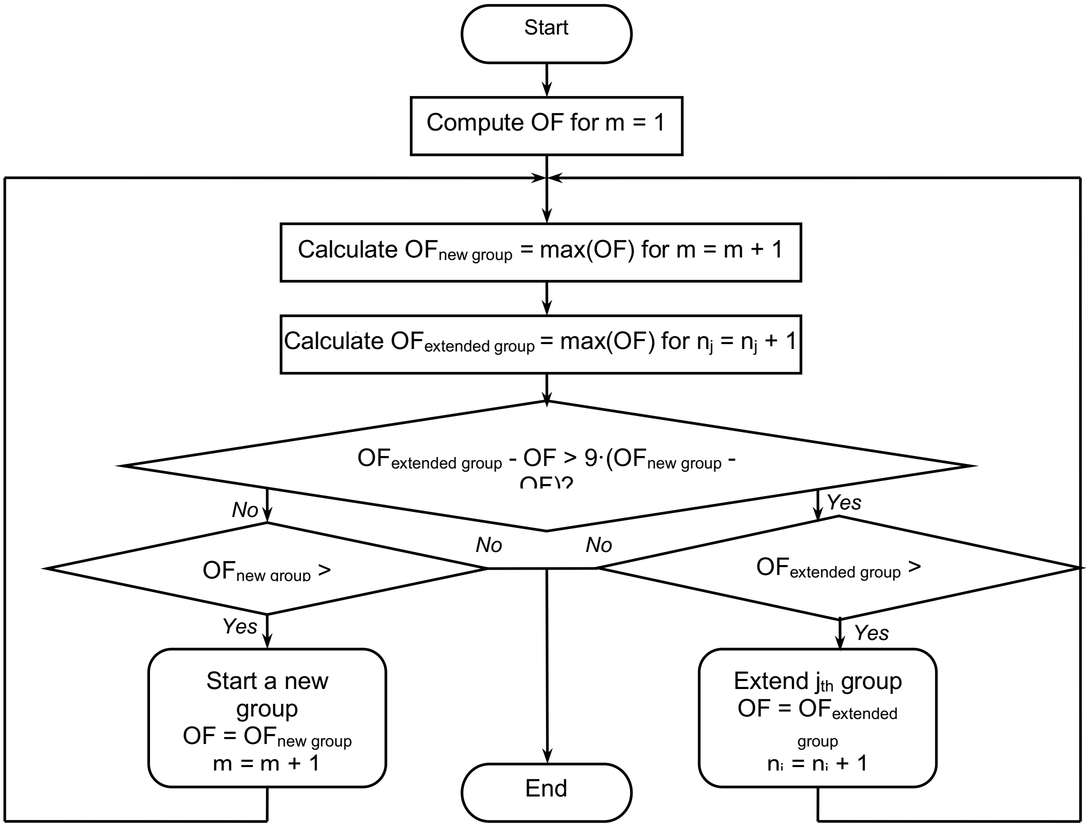
Start
Compute OF for m = 1
Calculate $OF_{new\ group}$ = max(OF) for m = m + 1
Calculate $OF_{extended\ group}$ = max(OF) for $n_j = n_j + 1$
$OF_{extended\ group}$ - OF > 9·($OF_{new\ group}$ - OF)?
No
Yes
$OF_{new\ group}$ >
No
No
$OF_{extended\ group}$ >
Yes
Yes
Start a new group
OF = $OF_{new\ group}$
m = m + 1
End
Extend j_{th} group
OF = $OF_{extended\ group}$
$n_i = n_i + 1$

Fig 2. The flowchart of the NSA algorithm process steps.

Interpolated maps of selected sensor variables

Kriged maps (spatial resolution of 5m) of RTK elevation (DEM), derived variables (TWI and slope), and Dualem sensor variables (HCP1, HCP2, PRP1, and PRP2) were produced and extracted in a data file for an input into the NSA tool. NDVI maps (spatial resolution of 10 m) were produced from Sentinel-2 images of 2017. Those continuous maps represented significant variations in different parts of each field (Figs. 3, 4, and 5).

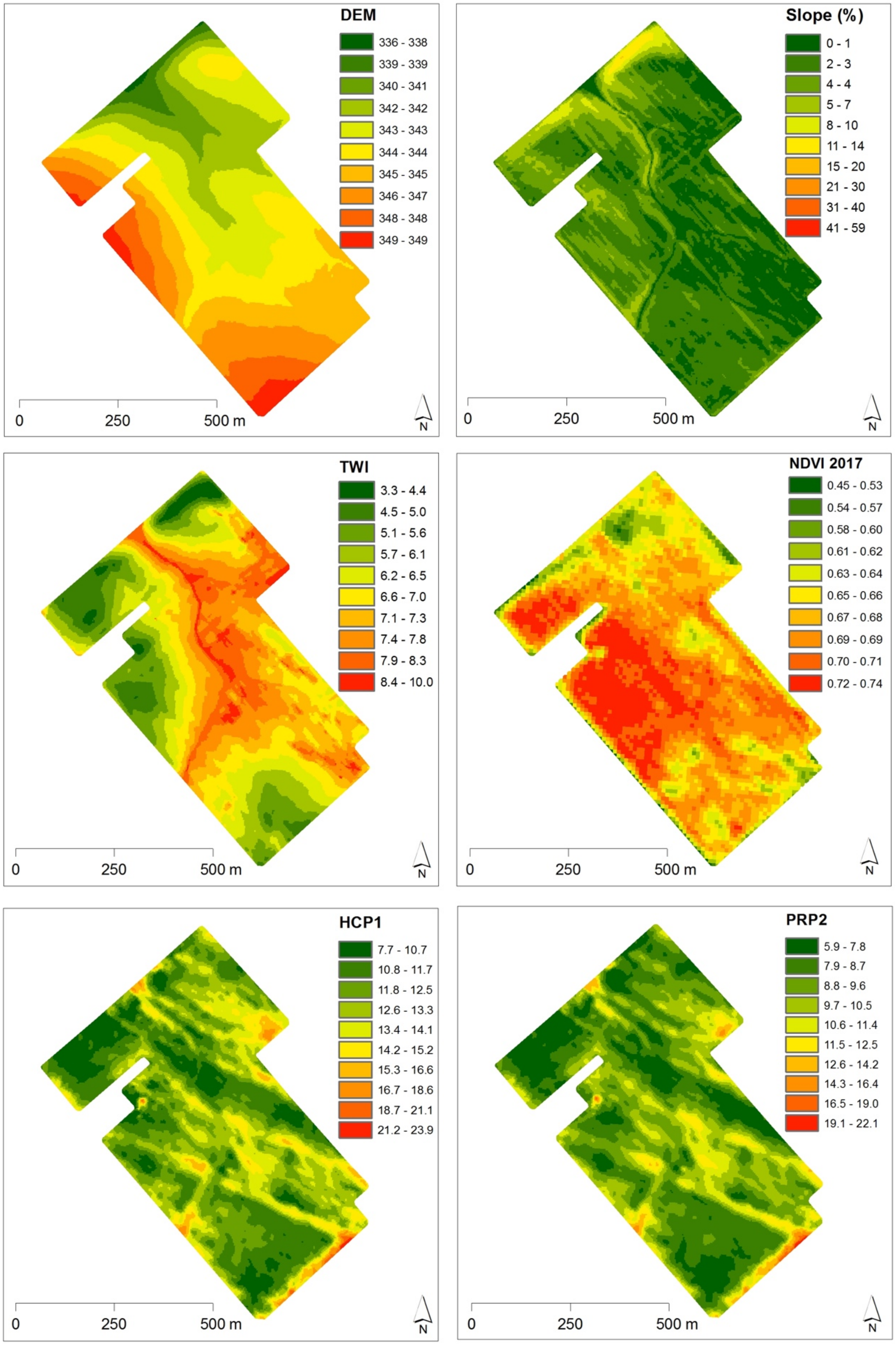

Fig 3. Interpolated maps (Kriged) of DEM, TWI, HCP1, PRP2 and NDVI maps in WH field.

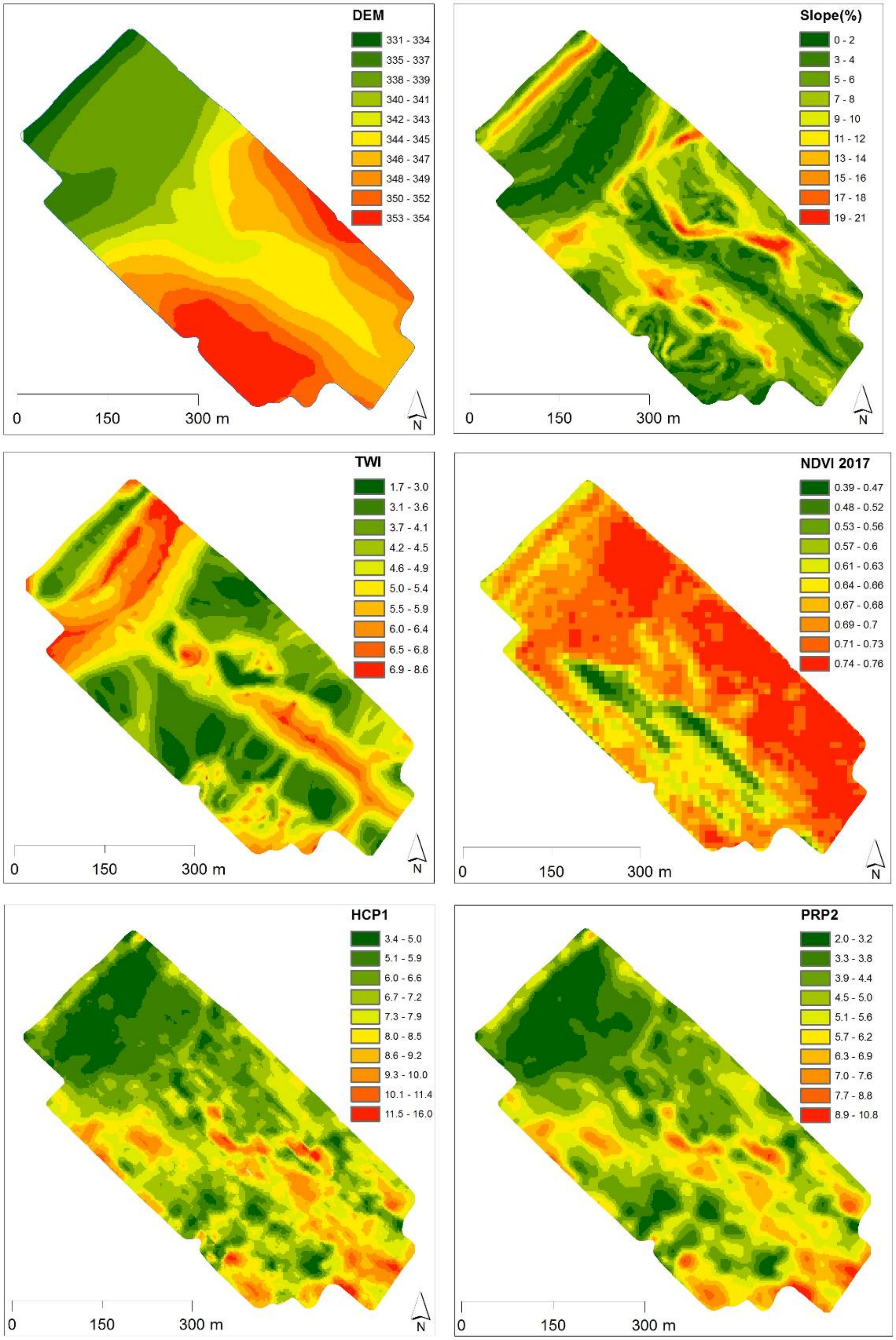

Fig 4. Interpolated maps (Kriged) of DEM, TWI, HCP1, PRP2 and NDVI maps in LD field.

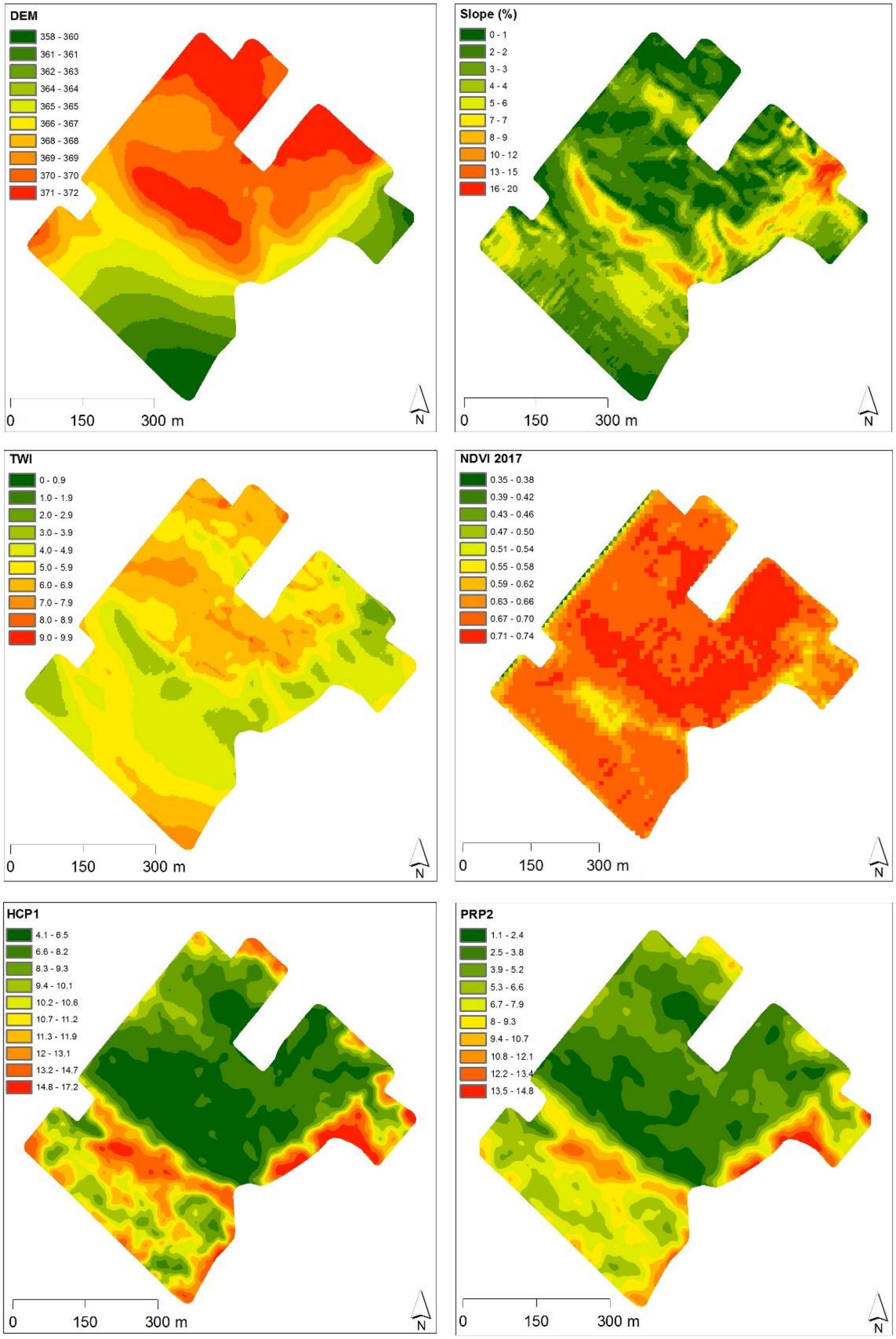

Fig 5. Interpolated maps (Kriged) of DEM, TWI, HCP1, PRP2 and NDVI maps in RB field.

Results and Discussion

FCM clustering

Based on the seven input variables (i.e., elevation, TWI, slope, HCP1, HCP2, PRP1, and PRP2) of the WH field, NCE and FPI indices in FCM clustering were assessed for their performance in creating the optimum number of zones. The NCE index was compared to FPI which showed that the maximum value was reached only in zones 4 and 5 (Fig 6). This clustering method presents a flaw when it comes to obtaining the optimum number of zones (Albornoz et al., 2018). The FCM clusters produced pixels with isolated boundaries in various parts of the field (Nazeer & Sebastian, 2009). Many studies have reported this representation problem regarding the clustering of data due to the fuzzy boundary (Bragato, 2004; LI et al., 2007; Panda et al., 2012). In this method, user-defined numbers of clusters were produced without considering the geospatial locations of the dataset or their distances.

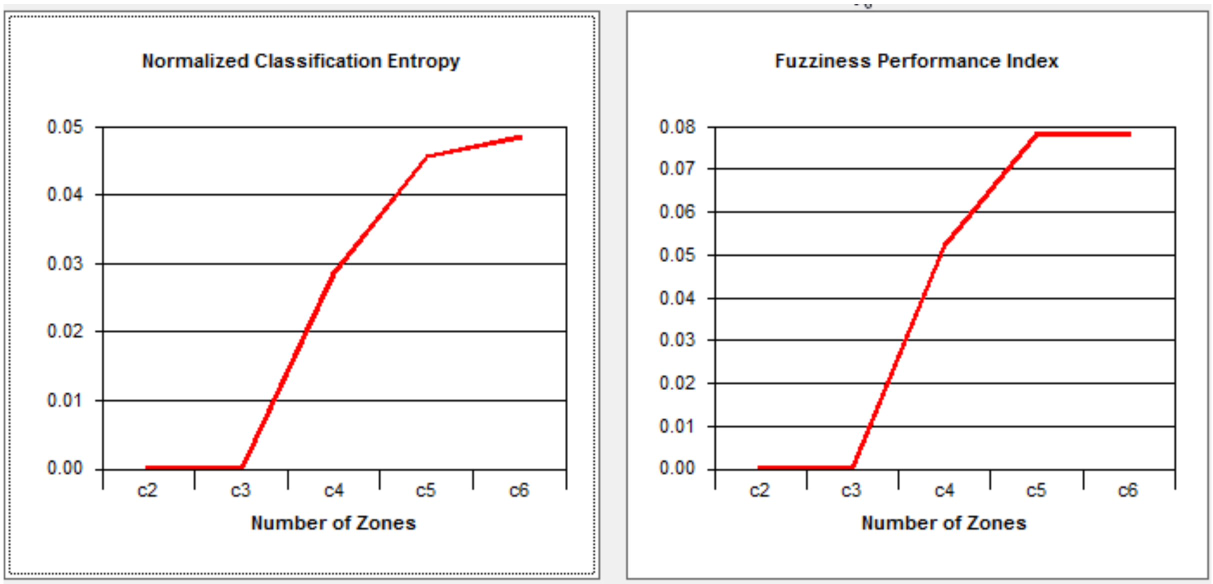

Fig 6. NCE and FPI of the WH field data with seven variables.

K-means clustering

In the K-means clustering (*K*=5), the data values were taken directly from the input table of WH field for generating cluster centers (Fig 7a). Data were standardized and normalized for the specific variable values. Among the five user-defined clusters, cluster 1, 2, 3, and 5 used the most data points. After several runs of each clustering process (K=5, K=15, and K=25), the R^2 were varied depending on how the K-means algorithm was initialized since there was a random component. The cluster map consisted of groups or pixels with isolated boundaries in various parts of the WH field (Fig 7b). Fig 7b, K-means cluster (K=25) map of WH field was produced for comparison with NSA zone map of approximately 25 clusters.

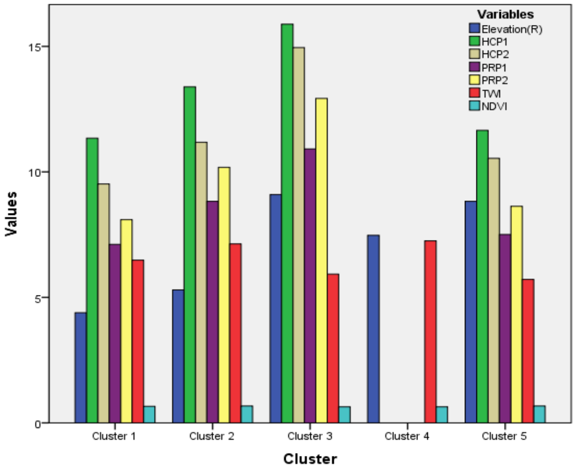

Fig 7(a). K-means cluster (*K*=5) centers with variable values of WH field.

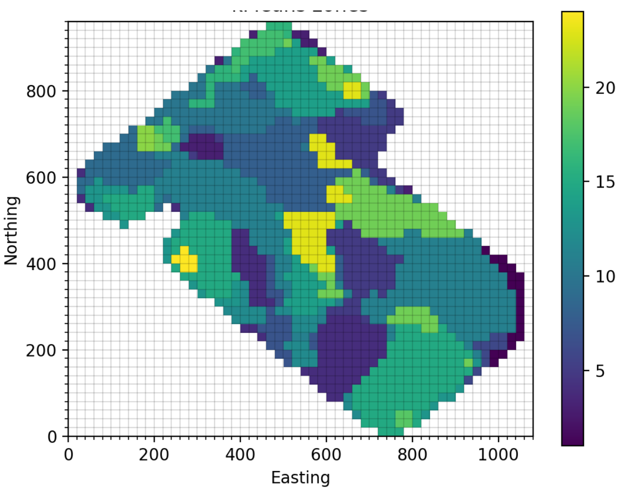

Fig 7(b). K-means cluster (*K*=25) map of WH field showing zones with various isolated pixels.

NSA clustering

In the NSA zone delineation process, providing the number of field partitioning clusters as compared to all other clustering algorithms is not obligatory. NSA produced groups of the grid cell (grid size of 20m) of seven input variables separately. This also could be delineated into user-defined zones. More importantly, this clustering tool efficiently delimited maps with the significant number of zones (Figs. 8a, 9a, and 10a). Results showed that WH, LD, and RB fields have 28, 20, and 27 georeferenced zones, respectively.

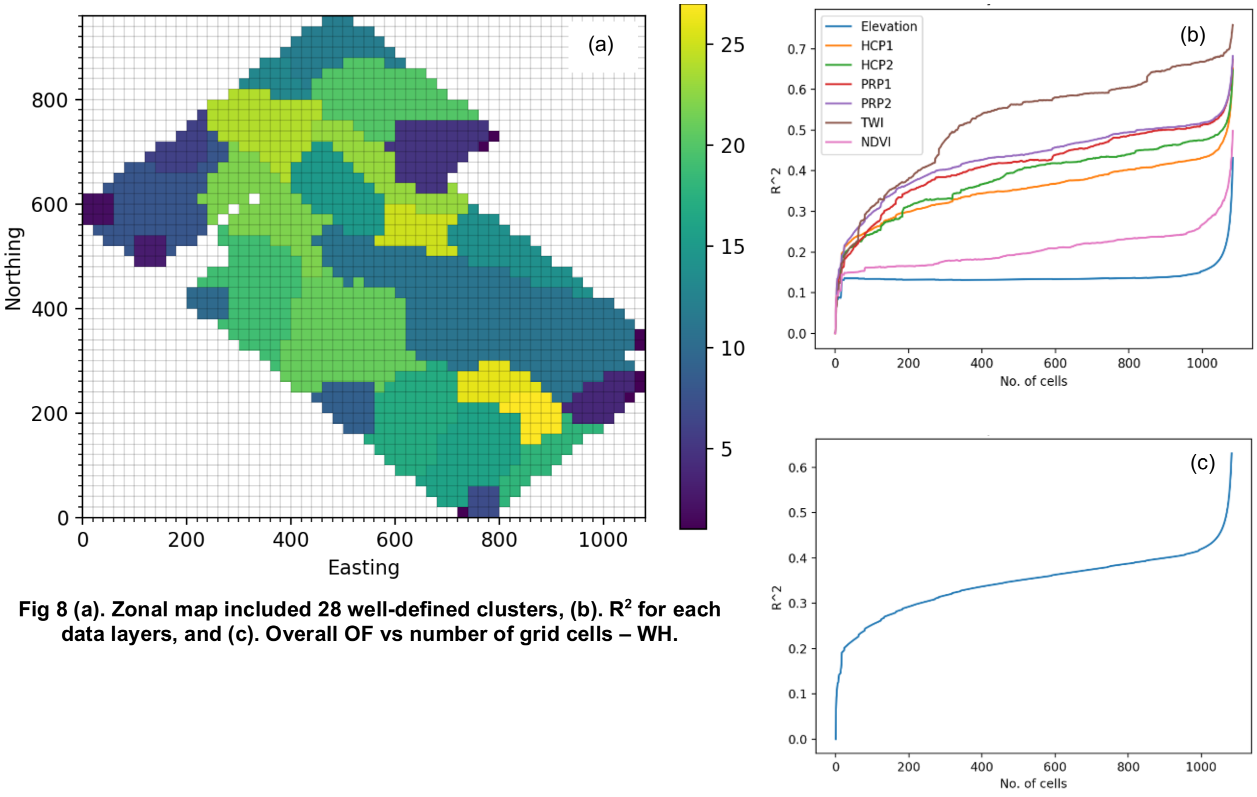

Fig 8 (a). Zonal map included 28 well-defined clusters, (b). R^2 for each data layers, and (c). Overall OF vs number of grid cells – WH.

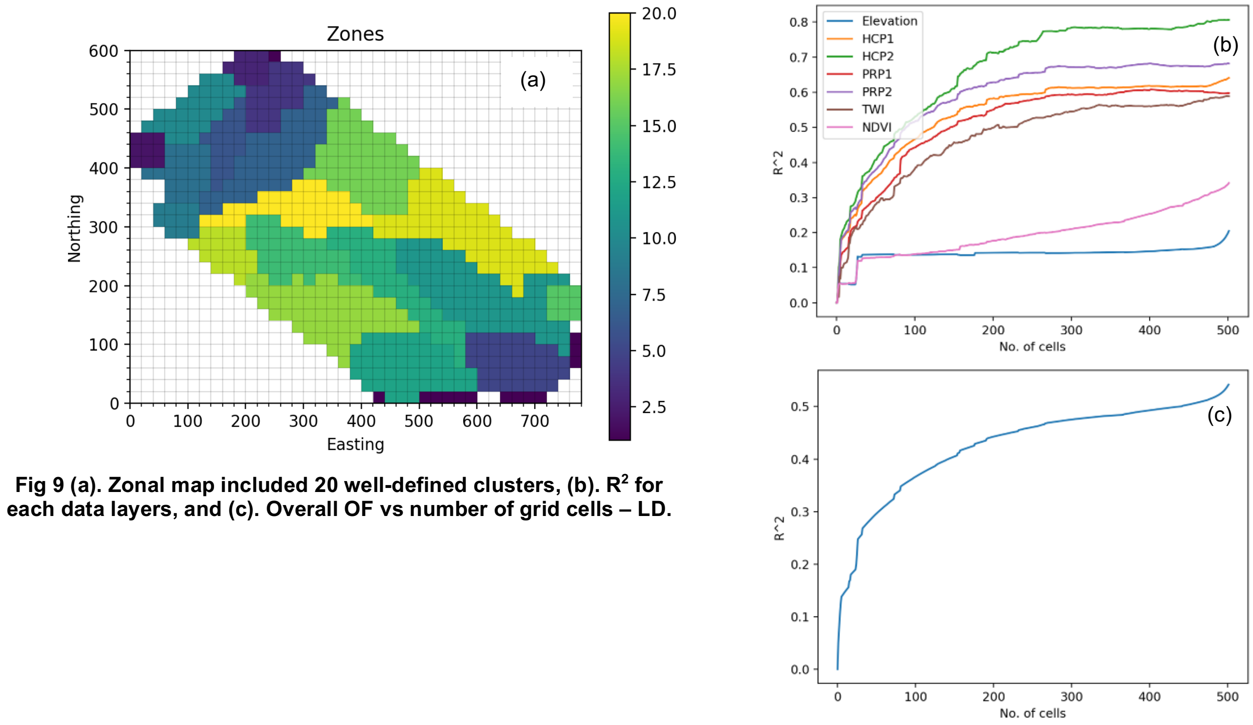

Fig 9 (a). Zonal map included 20 well-defined clusters, (b). R^2 for each data layers, and (c). Overall OF vs number of grid cells – LD.

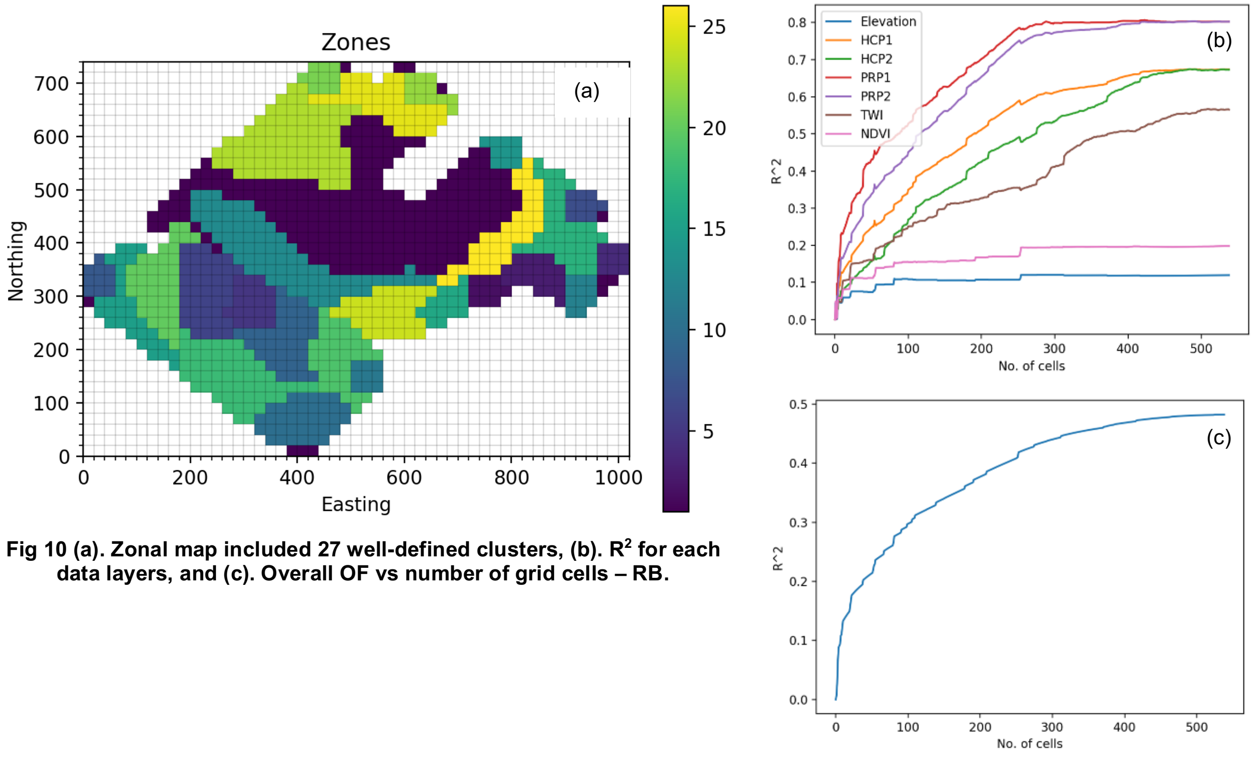

Fig 10 (a). Zonal map included 27 well-defined clusters, (b). R^2 for each data layers, and (c). Overall OF vs number of grid cells – RB.

Zone delineation was performed by the individual R^2 values of each variable (Figs. 8b, 9b, and 10b) and overall OF (Figs. 8c, 9c, and 10c). Those graphs showed which part of variance of each data layer was accounted for by subdividing the field into smaller areas. In each graph, a higher number of R^2 means that variability within individual zones was smaller than the difference between zones. Figs. 8b, 9b, and 10b showed that the R^2 values increased when any new groups were formed or added to the existing groups. The NSA produced R^2_{max} value was about 0.9 and the graph has a steeper slope in the beginning. This indicated that the data layer had a strong spatial structure and was dominated when the field was split. Also, x value, where most graphs level off, showed the smallest level of field partitioning that revealed majority of soil heterogeneity. Results in LD and RB fields indicated that R^2 for each data layer reached maximum height (0.5) around 400 classified grid cells, whereas the value reached at 0.65 near the 1000 grid cells in WH (Fig 11a). 50% (in LD and RB) and 65% (WH) of the field variance in both cases are accounted for by making the clusters.

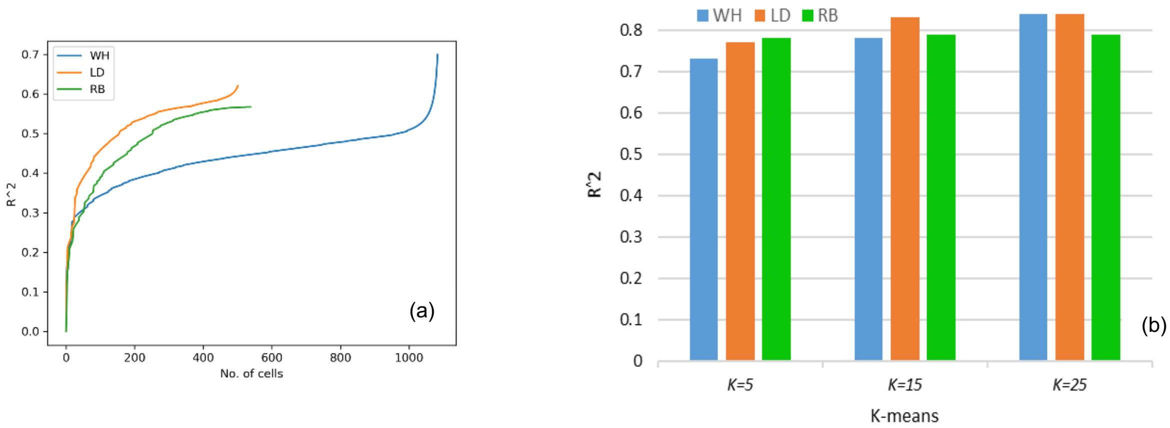

Fig 11. Comparison of R^2 value between K-Means and NSA clustering, (a). R^2 of NSA and (b). R^2 of K-means (K=5, K=15, and K=25).

R^2 values of NSA algorithm were compared among the three different fields (Fig 11a). The overall OF showed that all the clusters were maximized by the R^2 values closed to 0.5 and reached up to 0.65. In the three defined K-means clusters (K=5, K=15, and K=25), the R^2 of LD fields were reported higher: 0.78, 0.81, and 0.82 respectively (Fig 11b). Also, R^2 (K=5) was unexpectedly higher because of clumping pixels of each cluster throughout the field, and each cluster in various locations was not broken into parts. R^2 of K-means cluster compared to NSA was higher in most of the fields and was approximately 0.80. The R^2 values were comparable when the isolated/boundary pixels in each K-means cluster were disjointed from the main cluster. The K-means cluster map consisted of groups or pixels with isolated boundaries in various parts of the WH field (Fig 7b), whereas NSA algorithm counted these as different groups (Fig 8a).

Conclusions

The preprocessing and variable selection steps for all clustering techniques are imperative for providing a well-defined zonal boundary for developing management zones. Compared to other data clustering algorithms, NSA has a unique zone separation capability to produce a number of user-defined zones. Also, improved version of this sofware has been tested, which was handled a significant number of variables and data layers for delineating the optimum number of zones in a more robust way. The software was found reliable when integrating high-density field topography and PSS data files with the least amount of processing time, and it could be run on any platform with open source python modules. The robust zone delineation process and georeferenced thematic maps are useful for future applications of variable rate technologies and for other management purposes. In future, multisensor data fusion, advanced data filtering procedures, and the web application of the NSA could be implemented to help make appropriate site-specific agronomic and other environmental decisions in many regions.

Acknowlegements

This study was funded by the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) New Directions Research Program (NDRP). Special thanks to Doug Aspinall and to the Woodrill Farms for the data support and cooperation.

References

Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture*, 44(1), 71–91.

Adamchuk, V. I., Viscarra Rossel, R. A., Marx, D. B., & Samal, A. K. (2011). Using targeted sampling to process multivariate soil sensing data. *Geoderma*, *163*(1–2), 63–73.

Albornoz, E. M., Kemerer, A. C., Galarza, R., Mastaglia, N., Melchiori, R., & Martínez, C. E. (2018). Development and evaluation of an automatic software for management zone delineation. *Precision Agriculture*, *19*, 463–476.

Bragato, G. (2004). Fuzzy continuous classification and spatial interpolation in conventional soil survey for soil mapping of the lower Piave plain. *Geoderma*, 118(1–2), 1–16.

Burrough, P. A., Van Gaans, P. F. M., & Hootsmans, R. (1997). Continuous classification in soil survey: Spatial correlation, confusion and boundaries. *Geoderma*, 77(2–4), 115–135.

Cohen, S., Cohen, Y., Alchanatis, V., & Levi, O. (2013). Combining spectral and spatial information from aerial hyperspectral images for delineating homogenous management zones. *Biosystems Engineering*, *114*(4), 435–443.

Córdoba, M. A., Bruno, C. I., Costa, J. L., Peralta, N. R., & Balzarini, M. G. (2016). Protocol for multivariate homogeneous zone delineation in precision agriculture. *Biosystems Engineering*, 143, 95–107.

Cressie, N., & Kang, E. L. (2010). High-resolution digital soil mapping: Kriging for very large datasets. In *Proximal Soil Sensing* (pp. 49–63). Springer.

De Benedetto, D., Castrignano, A., Diacono, M., Rinaldi, M., Ruggieri, S., & Tamborrino, R. (2013). Field partition by proximal and remote sensing data fusion. *Biosystems Engineering*, 114(4), 372–383.

Dhawale N., Adamchuk V., Huang H., Ji W., Lauzon S., Biswas A., D. P. (2016). Integrated Analysis of Multilayer Proximal Soil Sensing Data. In International Conference on Precision Agriculture. St. Louis, Missouri, USA.

Dhawale, N. M., Adamchuk, V. I., Prasher, S. O., Dutilleul, P. R. L., & Ferguson, R. B. (2014). Spatially constrained geospatial data clustering for multilayer sensor-based measurements. In *Geospatial Theory, Processing, Modeling and Applications* (Vol. 40, pp. 187–190).

Fridgen, J. J., Kitchen, N. R., Sudduth, K. a, Drummond, S. T., Wiebold, W. J., & Fraisse, C. W. (2004). Management Zone Analyst (MZA): Software for subfield management zone delineation. *Agronomy Journal*, *96*(1), 100–108.

Gui-Fen, C., Li-Ying, C., Guo-Wei, W., Bao-Cheng, W., Da-You, L., & Sheng-Sheng, W. (2007). Application of a spatial fuzzy clustering algorithm in precision fertilisation. *New Zealand Journal of Agricultural Research*, 50(5), 1249–1254.

LI, Y., Shi, Z., & Li, F. (2007). Delineation of Site-Specific Management Zones Based on Temporal and Spatial Variability of Soil Electrical Conductivity. *Pedosphere*, 17(2), 156–164.

Nazeer, K. A. A., & Sebastian, M. P. (2009). Improving the Accuracy and Efficiency of the k-means Clustering Algorithm. In *Proceedings of the World Congress on Engineering* (Vol. I)1-5.

Panda, S., Sahu, S., Jena, P., & Chattopadhyay, S. (2012). Comparing fuzzy-C means and K-means clustering techniques: A comprehensive study. *Advances in Intelligent and Soft Computing*, 166(1), 451-460.

Roberts, D. F., Adamchuk, V. I., Shanahan, J. F., Ferguson, R. B., & Schepers, J. S. (2011). Estimation of surface soil organic matter using a ground-based active sensor and aerial imagery. *Precision Agriculture*, *12*(1), 82–102.

Ruß, G., & Brenning, A. (2010). Data Mining in Precision Agriculture: Management of Spatial Information. In E. Hüllermeier, R. Kruse, & F. Hoffmann (Eds.), Computational Intelligence for Knowledge-Based Systems Design (pp. 350–359). Berlin, Heidelberg: Springer.

U.S. Department of Agriculture (USDA). (2000). *Management Zone Analyst Version 1.0 Software*. Agricultural Research Service & University of Missouri Columbia, USA.

Viña, A., Gitelson, A. A., Nguy-robertson, A. L., & Peng, Y. (2011). Remote Sensing of Environment Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment*, 115(12), 3468–3478.

Viscarra Rossel, R.A., Adamchuk, V.I., Sudduth, K.A., McKenzie, N.J., and Lobsey, C. (2011). Proximal soil sensing: an effective approach for soil measurements in space and time. In *Advances in Agronomy* (pp. 237–283).

Vrindts, E., Mouazen, A. M., Reyniers, M., Maertens, K., Maleki, M. R., Ramon, H., & De Baerdemaeker, J. (2005). Management zones based on correlation between soil compaction, yield and crop data. *Biosystems Engineering*, 92(4), 419–428.