APPLICATIONS OF NUMERICAL CLASSIFICATION FOR SOME SOILS OF AL-HASHIMIYA PROJECT IN BABIL PROVINCE

A. M. Saleh
Assistant Professor
Coll. of Agric.- University of Baghdad

ABSTRACT

Cluster analysis was applied in grouping soils based on their scores of the factors controlling soil variation within 718.326 ha area in Al-Hashimiya District that was located in Babil province. Forty five sites were randomly sampled and analyzed for hierarchical cluster analysis which has been used to group samples by using Ward's method and to develop soil maps. The spatial distribution of distinct groups of elements demonstrates the interplay of ECe, Gypsum content, and particle size distribution factors. Cluster analysis appears to be useful for revealing patterns of soil homogeneity and for identifying relationships among soil properties. Numerical analysis may be a helpful supplementary method for correlating soil surveys with large soil databases.

Key words: cluster analysis, hierarchical analysis, Al-Hashimiya District, dendrogram, numerical classification.

تطبيقات التصنيف العددي لبعض ترب مشروع الهاشمية–محافظة بابل

آمال محمد صالح
أستاذ مساعد
كلية الزراعة–جامعة بغداد
mhmmd_ml@yahoo.com

المستخلص

من الطرق الاحصائية المهمة والمستخدمة في التصنيف العددي هي أسلوب التحليل العنقودي والذي يعتمد بدوره على تحليل متغيرات محددة تعتمد على نقاط التشابه والاختلاف بين المفردات. أستخدم التحليل الهرمي التجميعي في تصنيف بعض مفردات الترب الواقعة ضمن منطقة الهاشمية / محافظة بابل والبالغة مساحتها 718.326 هكتار ، إذ جرى أنتخاب خمس وأربعون موقعا بطريقة عشوائية لغرض تصنيف مفردات الترب تصنيفا عدديا بأتباع طريقة Ward وأعداد خرائط الترب لتعكس نمط التوزيع المكاني لمفردات تلك الترب ، لذا فقد أختيرت صفات الترب التي تخدم غرض التصنيف: الأيصالية الكهربائية ، محتوى الجبس ، والتوزيع النسبي لحجوم دقائق التربة. وكان التحليل العنقودي من أفضل طرائق التصنيف العددي والمستخدم لغرض تجميع العناصر أو مفردات الترب في مجموعات بحيث تكون العناصر أو المفردات متجانسة داخل كل مجموعة ومختلفة عن المجموعات الأخرى وأمكانية أستخدامها كقاعدة بيانات في مجال مسوحات الترب.

كلمات مفتاحية: التحليل العنقودي، التحليل الهرمي، منطقة الهاشمية، البناء الهرمي التجميعي، التصنيف العددي.

***Received:5/6/2018, Accepted:27/9/2018**

INTRODUCTION

The principal aim of cluster analysis is to partition observations into a number of groups. A good outcome of cluster analysis will result in a number of clusters where the observations within a cluster are as similar as possible while the differences between the clusters are as large as possible. Cluster analysis must thus determine the number of classes as well as the memberships of the observations to the groups. To determine the group membership most clustering methods use a measure of similarity between the observations (24). The similarity is usually expressed by distances between the observations. cluster analysis can be applied as an "exploratory data analysis tool" to better understand the multivariate behaviour of a data set. It can, however, never be a "statistical proof" of a certain relationship between the variables or observations (28). While factor analysis uses the correlation matrix for extracting common "factors" from a given data set most cluster analysis techniques use distance measures to assign observations to a number of groups (6). Correlation coefficients lie between –1 and +1, with 0 indicating linear independence (8). Similarities were calculated between soil profiles, represented as shaded similarity matrices, these were transformed to distances, which allowed a representation of the multi dimensional space in a few dimensions to be calculated. This is called ordination (11). Cluster analysis is a general term for a family of statistical classification methods that group objects. The idea is statistically to minimize within-group variability while maximizing among-group variability in order to produce relatively homogeneous groups that are distinct from one another (24). Cluster analysis has been used to develop conceptual schemes for grouping soils. Martín et al. (12) used the similarities among particle-size distributions to cluster soils, showing that the cluster classes approximated existing series. The indices form a matrix are used to construct a dendrogram which illustrates the clustering into groups and good agreement with field observations was obtained. Data analysis in soil classification studies is easier with statistical tools such as factor analysis and hierarchical classification which are explanatory techniques. Cluster analysis and Pearson's correlation matrix have proven to be useful in offering reliable classification of the metals and physicochemical properties of soils (10). Agglomerative hierarchical clustering starts from a proximity matrix between individuals, each one forming a singleton cluster, and gathers clusters into groups of clusters or superclusters, the process being repeated until a complete hierarchy of partitions into clusters is formed (25). Since numerical classification of soils is impeded by the so-called anisotropy of the profiles (not all horizons occur everywhere) (30), most authors by passed this problem using soil samples taken at fixed depths (20). Furthermore, most of the numerical classification studies used quantitative chemical and physical soil properties, which makes them difficult to apply in the field (21). Therefore, the aim of this study is to explore the possibilities of a numerical soil classification system which starts from soil properties and uses soil horizons towards soil profile membership classification. With this classification system an attempt is made at Al-Hashimiya, Babil to produce high-resolution soil classes, which remain compatible to existing higher order frameworks for soil classification.

MATERIALS AND METHODS

This study was conducted within Al-Hashimiya District is a part of Babil province, 100 km (62 mi) south of Baghdad, rising 34 m above sea level. The study area was about 718.326 ha which was located between 44° 27' 23.48" to 44° 27' 38.25" of Eastern longitude and 32° 21' 14.531" to 32° 25' 5.337" of Northern latitude (Figs. 1, 2). The climate is considered to be BWh according to the Köppen-Geiger climate classification (22). Temperatures can reach as high as 50 °C, the average annual temperature is 23.1 °C, and winters are generally mild. The rainfall here averages 114 mm (Iraqi Meteorology, 2016). Soil samples were randomly taken from 45 locations in March 2016 (Al-Hashimiya Project, 2016). The locations of sampling sites were identified by global positioning system (GPS) showed in Fig. (3). Soil samples were taken from A_1, C_1, and C_2 horizons of soil

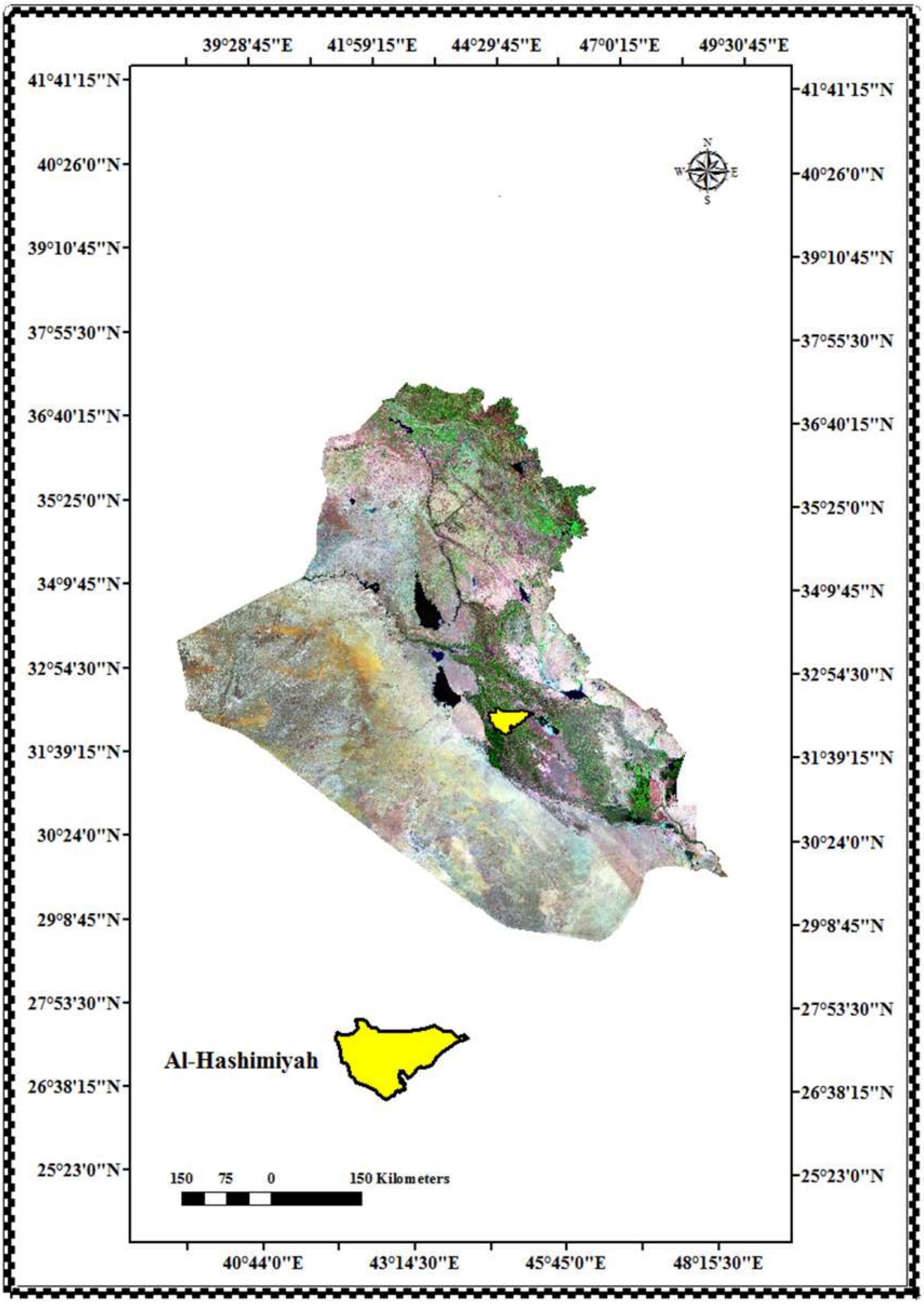

Figure 1. Location of the study area at Al-Hashimiya District, Babil, Iraq.

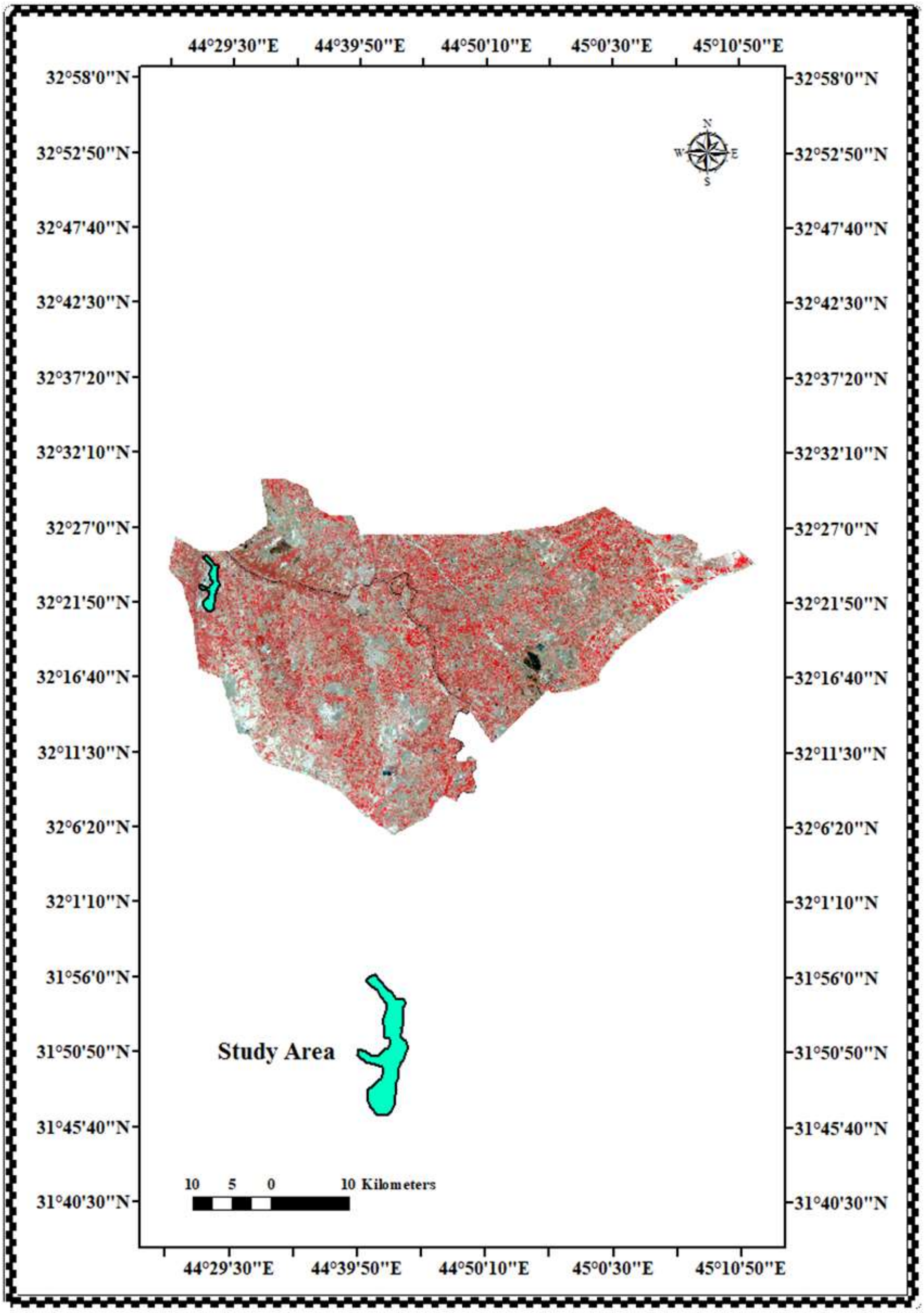

Figure 2. Location of Al-Hashimiya Project, Babil, Iraq.

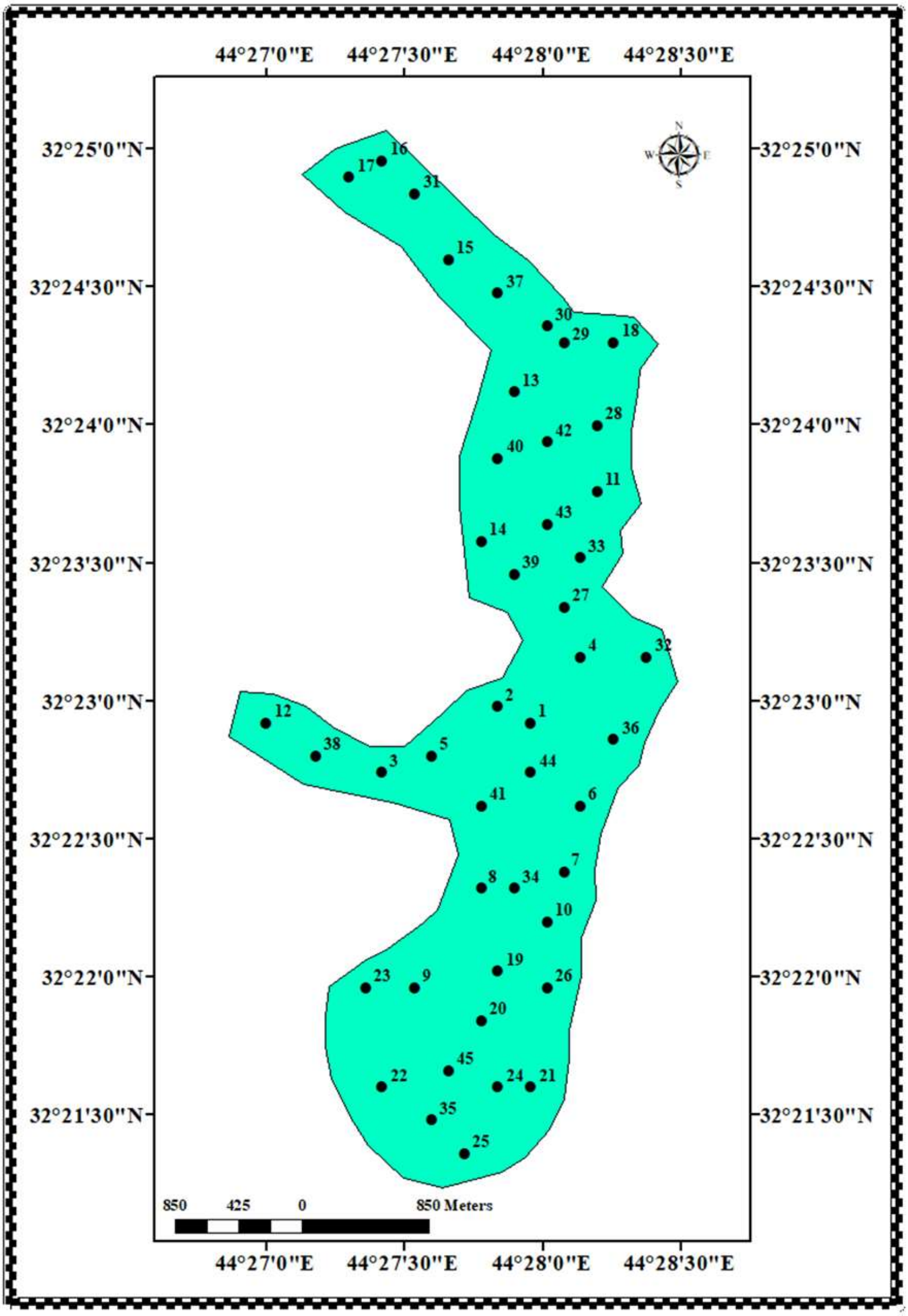

Figure 3. Map of the study area showing 45 soil locations.

profiles, and analyzed for particle size distribution, electrical conductivity (ECe), and Gypsum content ($CaSO_4$). The data were collected from the Ministry of Water Resources/National Center for Water Resources Management for characterization of the standard physical and chemical properties of the soil samples at the study area. Based on morphological properties and physicochemical analysis, soil individuals were classified as Entisols and Aridisols according to Soil Survey Staff (26), (27) (Table 1). Statistical analysis was performed with Statistical Programme for Social Sciences (SPSS® software) version 16.0 for the computation of the hierarchical cluster analysis, which represents a quantitative independent approach of soil individuals and variables classification in environmental studies. Hierarchical cluster analysis was performed to identify analogous behaviour among the different characteristics of soils and also among soil individuals. It was performed on the normalized data set by means of Ward's method using squared Euclidean distances as a measure of similarity between soil individuals (19). There are two approaches to hierarchical clustering: we can go\from the bottom up", grouping small clusters into larger ones, or \from the top down",splitting big clusters into small ones. These are called agglomerative and divisive clusterings, respectively. In this study we stick to agglomerative clustering. Ward's method says that the distance between two clusters, A and B, is how much the sum of squares will increase when we merge them:

$$\Delta(\boldsymbol{A},\boldsymbol{B}) = \sum_{i\in A\cup B}\|\vec{x}_i - \vec{m}_{A\cup B}\|^2 - \sum_{i\epsilon A}\|\vec{x}_i - \vec{m}_A\|^2 - \sum_{i\in B}\|\vec{x}_i - \vec{m}_B\|^2 \quad \text{.......(1)}$$

$$= \frac{n_A n_B}{n_A + n_B}\|\vec{m}_A - \vec{m}_B\|^2 \quad \text{.......(2)}$$

where $\vec{m}_j$ is the center of cluster j and n_j is the number of points in it. Δ is called the merging cost of combining the clusters A and B. With hierarchical clustering, the sum of squares starts out at zero (because every point is in its own cluster) and then grows as we merge clusters. Ward's method keeps this growth as small as possible (7). Minasny and McBratney (13) indicate that characters should be as numerous as possible, free of inter-influences, and that they should be treated with equal importance (i.e.with equal weights).Soil provide numerous characteristics for use in classification. However, the charactes are not independent of each other. Soil scientists are reluectant to give equal weights to all soil characters. Fifteen soil characteristics were used to arry 45 soil individuals in a taxonomic dendrogram (Table 1), (Table 2).

RESUTS AND DISCUSSION

The first step in the hierarchical clustering process is to look for the pair of samples that are the most similar, that are the closest in the sense of having the lowest dissimilarity,. These two samples are then joined in the first step of the dendrogram, or clustering tree (the vertical scale of 0 to 25 which calibrates the level of clustering). The point at which they are joined is called a node. Hierarchical agglomerative clustering HAC is a bottom-up technique to generate a tree-like structure of clusters called dendrogram. In each level of the dendrogram, a full clustering of the underlying data is depicted. HAC usually starts at the level of single data records and consecutively merges records and/or clusters, thereby creating the dendrogram. In the agriculture data, hierarchical clustering may be applied, given that two specialties are taken into account. First, due to spatial autocorrelation, the lowest level from which HAC starts may be replaced by contiguous zones which can be generated using a spatial tessellation. Second, HAC may proceed as usual but should consider a spatial constraint: since the resulting management zones are suggested to be contiguous, only spatially neighboring zones are to be merged (16). Tables (3), (4), and (5) showed the results of the first step with the assignment of 9, 9, and 10 provisional clusters, respectively. They summarize the number of samples, which fell into a respective cluster, the ranges and the means relevant to five variables used in the analysis. In the second step, hierarchical cluster analysis followed in order to group the provisional cluster into some final clusters. Results of dendrograms in Figs. (4), (5) and (6) showed that the all clusters were identified through two different trials, agreed well with each other at the 0.05 level (2-tailed) except in the "silt content" factor scores.

Table 1. Soil individuals of the study area.

Soil number	Order	Class of lower category	Location
1.	Entisols	Typic Torrifluvents	32.382^o N - 44.466^o E
2.	Aridisols	Gypsic Haplosalids	32.383^o N - 44.464^o E
3.	Entisols	Typic Torrifluvents	32.379^o N - 44.457^o E
4.	Entisols	Typic Torrifluvents	32.386^o N - 44.469^o E
5.	Entisols	Typic Torrifluvents	32.380^o N - 44.460^o E
6.	Entisols	Typic Torrifluvents	32.377^o N - 44.469^o E
7.	Entisols	Typic Torrifluvents	32.373^o N - 44.468^o E
8.	Entisols	Typic Torrifluvents	32.372^o N - 44.463^o E
9.	Aridisols	Gypsic Haplosalids	32.366^o N - 44.459^o E
10.	Entisols	Typic Torrifluvents	32.370^o N - 44.467^o E
11.	Entisols	Typic Torrifluvents	32.396^o N - 44.470^o E
12.	Entisols	Typic Torrifluvents	32.382^o N - 44.450^o E
13.	Aridisols	Typic Haplosalids	32.402^o N - 44.465^o E
14.	Aridisols	Typic Haplogypsids	32.393^o N - 44.463^o E
15.	Entisols	Typic Torrifluvents	32.410^o N - 44.461^o E
16.	Entisols	Typic Torrifluvents	32.416^o N - 44.457^o E
17.	Entisols	Typic Torrifluvents	32.415^o N - 44.455^o E
18.	Entisols	Typic Torrifluvents	32.405^o N - 44.471^o E
19.	Entisols	Typic Torrifluvents	32.367^o N - 44.464^o E
20.	Entisols	Typic Torrifluvents	32.364^o N - 44.463^o E
21.	Aridisols	Typic Haplogypsids	32.360^o N - 44.466^o E
22.	Entisols	Gypsic Haplosalids	32.360^o N - 44.457^o E
23.	Entisols	Gypsic Haplosalids	32.366^o N - 44.456^o E
24.	Entisols	Typic Torrifluvents	32.360^o N - 44.464^o E
25.	Entisols	Typic Torrifluvents	32.356^o N - 44.462^o E
26.	Aridisols	Typic Haplogypsids	32.366^o N - 44.467^o E
27.	Entisols	Typic Torrifluvents	32.389^o N - 44.468^o E
28.	Entisols	Typic Torrifluvents	32.400^o N - 44.470^o E
29.	Entisols	Typic Torrifluvents	32.405^o N - 44.468^o E
30.	Entisols	Typic Torrifluvents	32.406^o N - 44.467^o E
31.	Aridisols	Typic Haplogypsids	32.414^o N - 44.459^o E
32.	Entisols	Typic Torrifluvents	32.386^o N - 44.473^o E
33.	Aridisols	Typic Haplogypsids	32.392^o N - 44.469^o E
34.	Entisols	Typic Torrifluvents	32.372^o N - 44.465^o E
35.	Entisols	Typic Torrifluvents	32.364^o N - 44.463^o E
36.	Entisols	Typic Torrifluvents	32.381^o N - 44.471^o E
37.	Entisols	Typic Torrifluvents	32.408^o N - 44.464^o E
38.	Entisols	Typic Torrifluvents	32.380^o N - 44.453^o E
39.	Entisols	Typic Torrifluvents	32.391^o N - 44.465^o E
40.	Entisols	Typic Torrifluvents	32.398^o N - 44.464^o E
41.	Entisols	Typic Torrifluvents	32.377^o N - 44.463^o E
42.	Entisols	Typic Torrifluvents	32.399^o N - 44.467^o E
43.	Entisols	Typic Torrifluvents	32.394^o N - 44.467^o E
44.	Entisols	Typic Torrifluvents	32.379^o N - 44.466^o E
45.	Aridisols	Gypsic Haplosalids	32.361^o N - 44.461^o E

Table 2. Soil individuals characteristics.

1.	ECe of A_1 Horizon
2.	ECe of C_1 Horizon
3.	ECe of C_2 Horizon
4.	Content of Gypsum of A_1 Horizon
5.	Content of Gypsum of C_1 Horizon
6.	Content of Gypsum of C_2 Horizon
7.	Content of Sand fraction of A_1 Horizon
8.	Content of Sand fraction of C_1 Horizon
9.	Content of Sand fraction of C_2 Horizon
10.	Content of Silt fraction of A_1 Horizon
11.	Content of Silt fraction of C_1 Horizon
12.	Content of Silt fraction of C_2 Horizon
13.	Content of Clay fraction of A_1 Horizon
14.	Content of Clay fraction of C_1 Horizon
15.	Content of Clay fraction of C_2 Horizon

Table 3. Summary of hierarchical cluster analysis based on A_1 soil horizon characters.

Final cluster	No. of members	ECe dS m^{-1}		Gypsum g kg^{-1}		Clay g kg^{-1}		Silt g kg^{-1}		Sand g kg-^{1}	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
A	10	39.85	8.1-95	63.85	9-104	360.23	320.6-440	547.80	490-572	90.20	60-110
B	1	7.03	-----	93.00	-----	190.00	-----	662.00	-----	148.00	-----
C	1	11.00	-----	40.00	-----	100.00	-----	570.00	-----	330.00	-----
D	4	7.47	4.9-10.94	18.32	3-51.5	282.50	240-300	632.00	604-674	85.50	70-100
E	3	5.05	2.16-7.8	9.00	6.6-12.3	415.00	385-440	511.67	506-520	63.33	40-76
F	7	8.27	3.2-16.36	13.00	6.1-23.7	328.06	322.5-343.5	567.84	562.7-572	101.77	89.6-105.6
G	8	3.95	0.7-10.94	14.71	0.5-51.5	341.07	300-380	570.9	549-630	86.61	58-105.6
H	10	10.52	3.8-18.57	21.8-	6.1-34	325.40	303-350	563.30	530-575	109.62	83-150
I	1	45.00	-----	30.70	-----	322.50	-----	572.00	-----	103.00	-----

Table 4. Summary of hierarchical cluster analysis based on C_1 soil horizon characters.

Final cluster	No. of members	ECe dS m^{-1}		Gypsum g kg^{-1}		Clay g kg^{-1}		Silt g kg^{-1}		Sand g kg-^{1}	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
A	8	18.13	11.3-23.9	12.90	10-18.8	343.17	330-356	549.04	545-557	94.77	87.4-110.7
B	3	3.63	2.8-5.3	6.66	0.9-12.1	316.66	295-330	612.66	595-622	70.66	48-84
C	4	6.63	4.1-8.4	17.80	14-23	293.72	280-305	549.80	530-557.3	138.77	120-170
D	3	4.06	1.68-8.70	3.33	2-4.4	322.46	300-353.6	550.30	548-554.6	118.96	87-152
E	9	6.75	2.8-22.6	14.21	12.2-16	353.04	338.8-366	550.84	548-553.8	102.00	101.5-107
F	8	6.72	4.1-11	11.62	10-14.8	374.60	340.2-420	542.68	520-552.6	87.30	50-105.8
G	3	4.15	1.5-5.6	7.10	0.3-12.1	413.33	295-500	526.66	464-621	60.00	36-84
H	4	39.70	30.4-50.5	17.30	8-30.2	378..02	327-440	525.40	495-556	93.67	62-110
I	3	8.60	6.4-11.17	34.46	13-47.4	167.33	92-300	629.33	578-678	203.33	68-330

Cluster C in Fig. (4) was compared to cluster I in Fig. (5) and both of them showed a slightly gypsiferous value in terms of the "gypsum content" factor scores according to Barazanji (4), on the other hand, cluster C in Fig. (4) had no significant characteristics with the rest of the clusters at the 0.05 level (2-tailed) due to a lesser extent of "clay content". Hierarchical clustering of Al-Hashimiya soils showed that the elements were associated with salinity that could be interpreted in terms of parent material, surface-water salinity, Irrigation and drainage (3). Gypsiferous soils can readily be classified as highly sensitive to environmental conditions such as salts accumulation on the soil surface (1).

Table 5. Summary of hierarchical cluster analysis based on C_2 soil horizon characters.

Final cluster	No. of members	ECe dS m^{-1}		Gypsum g kg^{-1}		Clay g kg^{-1}		Silt g kg^{-1}		Sand g kg^{-1}	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
A	7	7.39	6.07-8.88	10.86	9.4-14.5	448.43	432.7-473.6	509.54	488-525.9	42.00	35-51
B	1	7.26	-----	10.80	-----	540.00	-----	442.00	-----	18.00	-----
C	6	3.94	1.28-5.60	7.11	0.1-11.4	395.70	365-420	546.50	525-565	62.40	52-79.8
D	3	20.35	17.6-24.3	5.36	6-4.1	382.76	358.3-410	545.03	520-563.1	72.33	70-79
E	4	15.35	5.6-21.8	11.10	8-15.8	320.52	296.4-354	585.52	569.5-598	94.00	66.3-110.2
F	1	31.1	-----	1.60	-----	340.00	-----	528.00	-----	132.00	-----
G	2	5.93	5.29-6.58	40.85	39-42.7	267.50	200-335	614.00	566-662	118.50	99-138
H	9	6.05	3.3-9.8	10.64	0.3-20	323.40	277.5-463.4	589.33	488-616.7	86.82	51-106.6
I	8	7.84	1.76-17.1	9.02	4.3-12.5	336.85	310.2-365	574.88	552-597.7	89.44	70-104.7
J	4	10.80	6.6-13.15	4.55	1-6.8	233.30	145-310	612.15	532-671	152.07	118.6-184

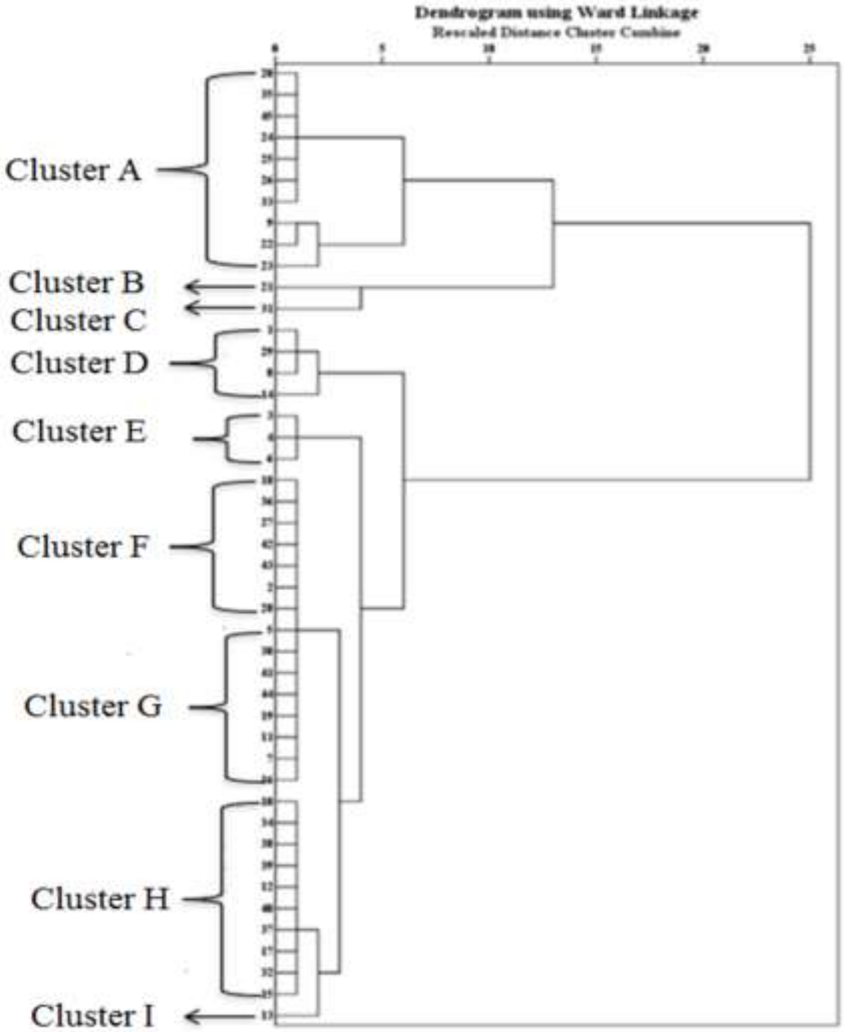

Figure 4. Taxonomic dendrogram based on A_1 soil horizon characters.

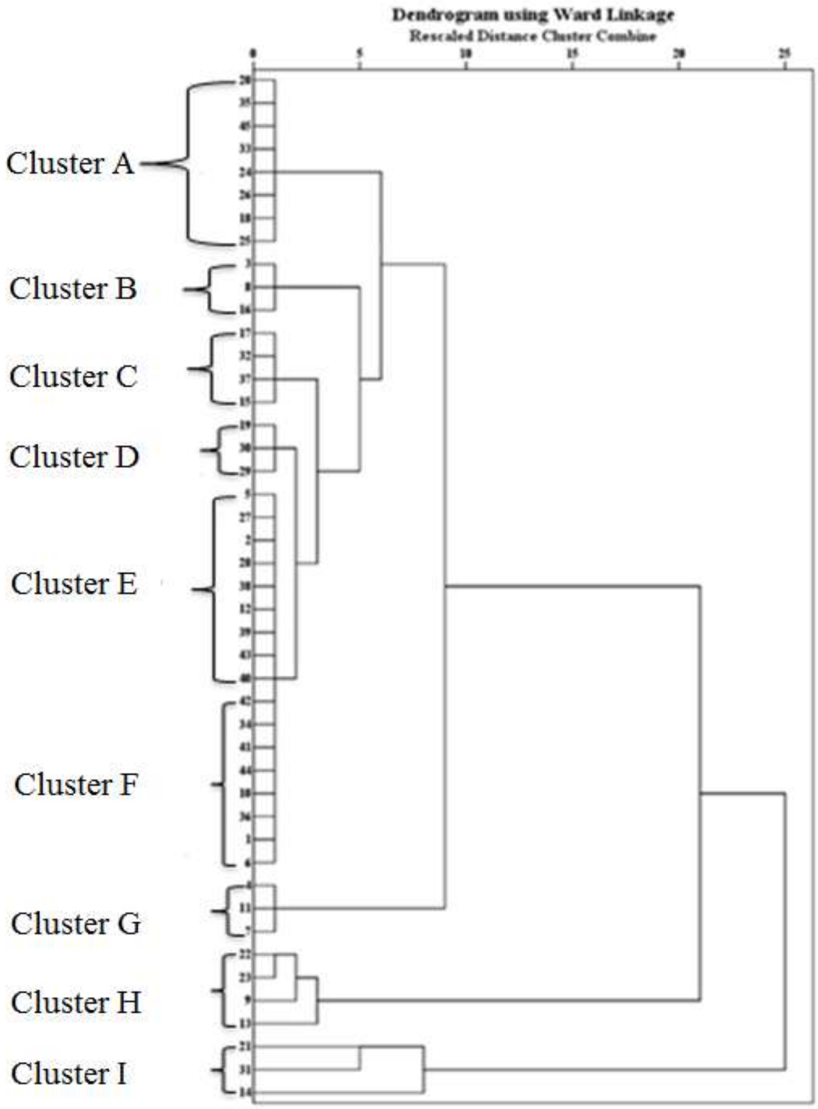

Figure 5. Taxonomic dendrogram based on C_1 soil horizon characters.

Cluster A of hierarchical clustering of A_1 soil horizon in Fig. (4) showed elements associations with extremely salinity (5) and slightly gypsiferous in the dendrogram. Elements of cluster F were typically associated with moderately salinity and very slightly gypsiferous. Also elements of cluster H were closely related in the dendrogram, further these elements were commonly associated with highly salinity and very slightly gypsiferous. Cluster I showed element distinctly separated from elements of cluster H based on extremely salinity spatial trends of the clusters (23). Elements of Cluster A of hierarchical clustering of C_1 soil horizon in Fig. (5) were strongly associated with highly salinity in the dendrogram. Elements of clusters E and F were also closely related in the dendrogram. However, they were associated with moderately salinity but cluster F separated from elements of cluster E due to a lesser extent of "sand content" factor scores. Clusters A and B of hierarchical clustering of C_2 soil horizon in Fig. (6) showed elements associations with moderately salinity and very slightly gypsiferous but cluster B separated from elements of cluster A due to a higher extent of "clay content" factor scores. Element of cluster F with severely salinity and non gypsiferous distinctly separated from elements of cluster E in the dendrogram. As described above, four outstanding clusters were recognized (Fig. 7). One of them was however, apparently a much larger cluster than the others. Although clusters with special soil characteristics were identified, it is considered that the procedure may have overlooked the minor differences. It may be more useful to assign multiple disjoint levels for the practical classification of the soils. The soil map shown in Fig. (8) was compiled based on A_1 soil horizon characters using Ward's method. Clusters A and B were located in the south - western corner of the study area. In general, these two clusters covered 9.52%-8.04% of the total study area respectively.

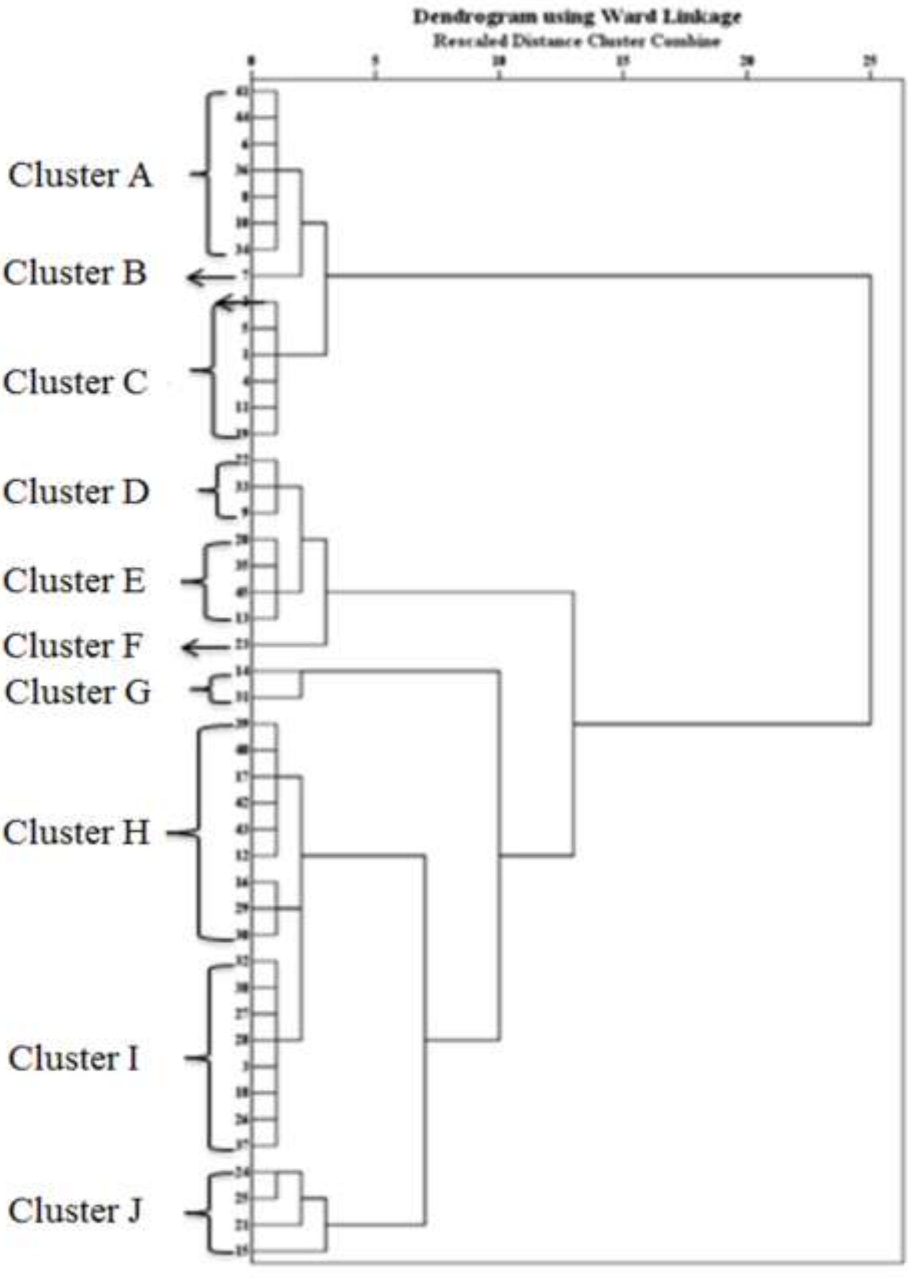

Figure 6. Taxonomic dendrogram based on C_2 soil horizon characters.

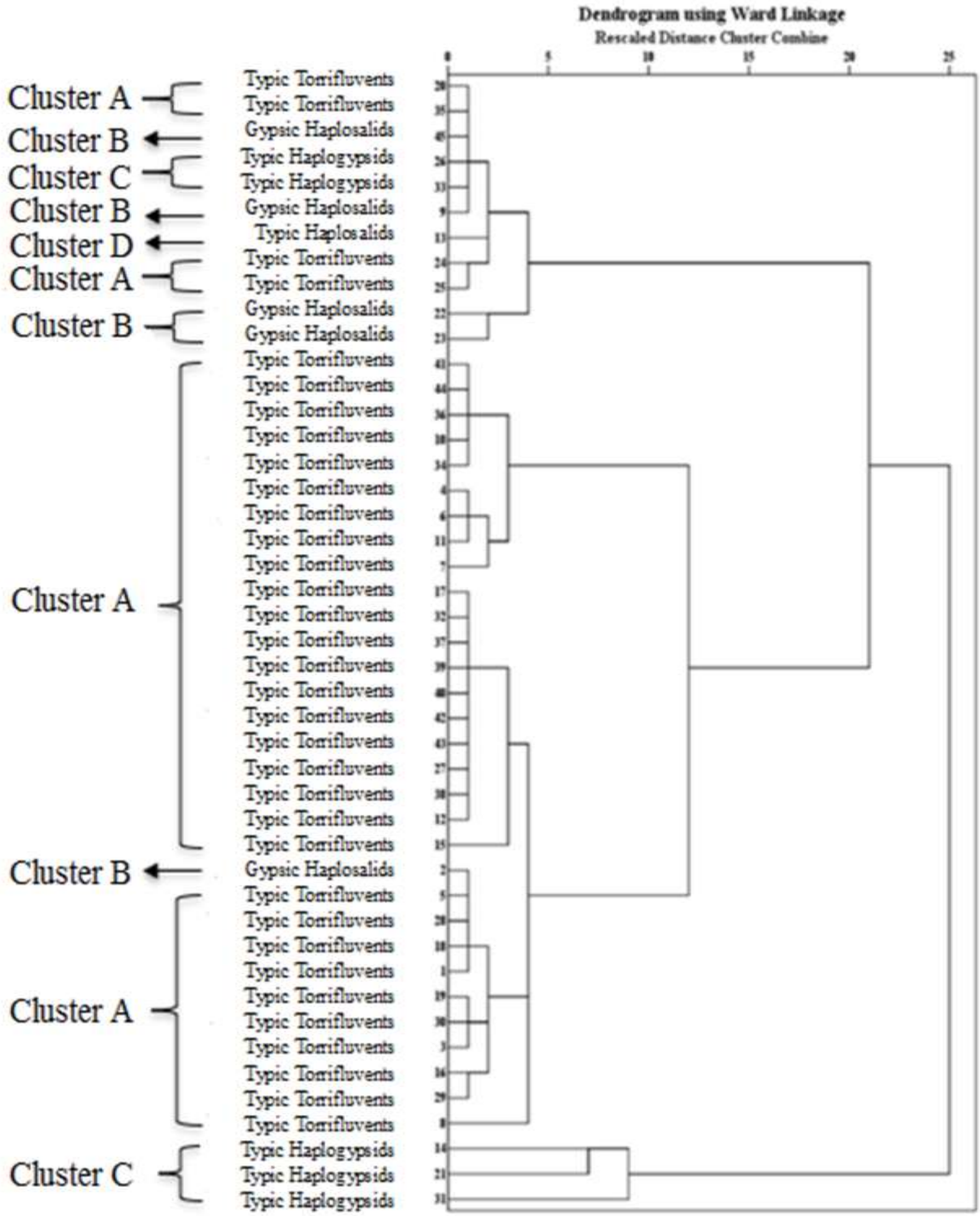

Figure 7. Taxonomic dendrogram based on A_1, C_1, and C_2 soil horizon characters.

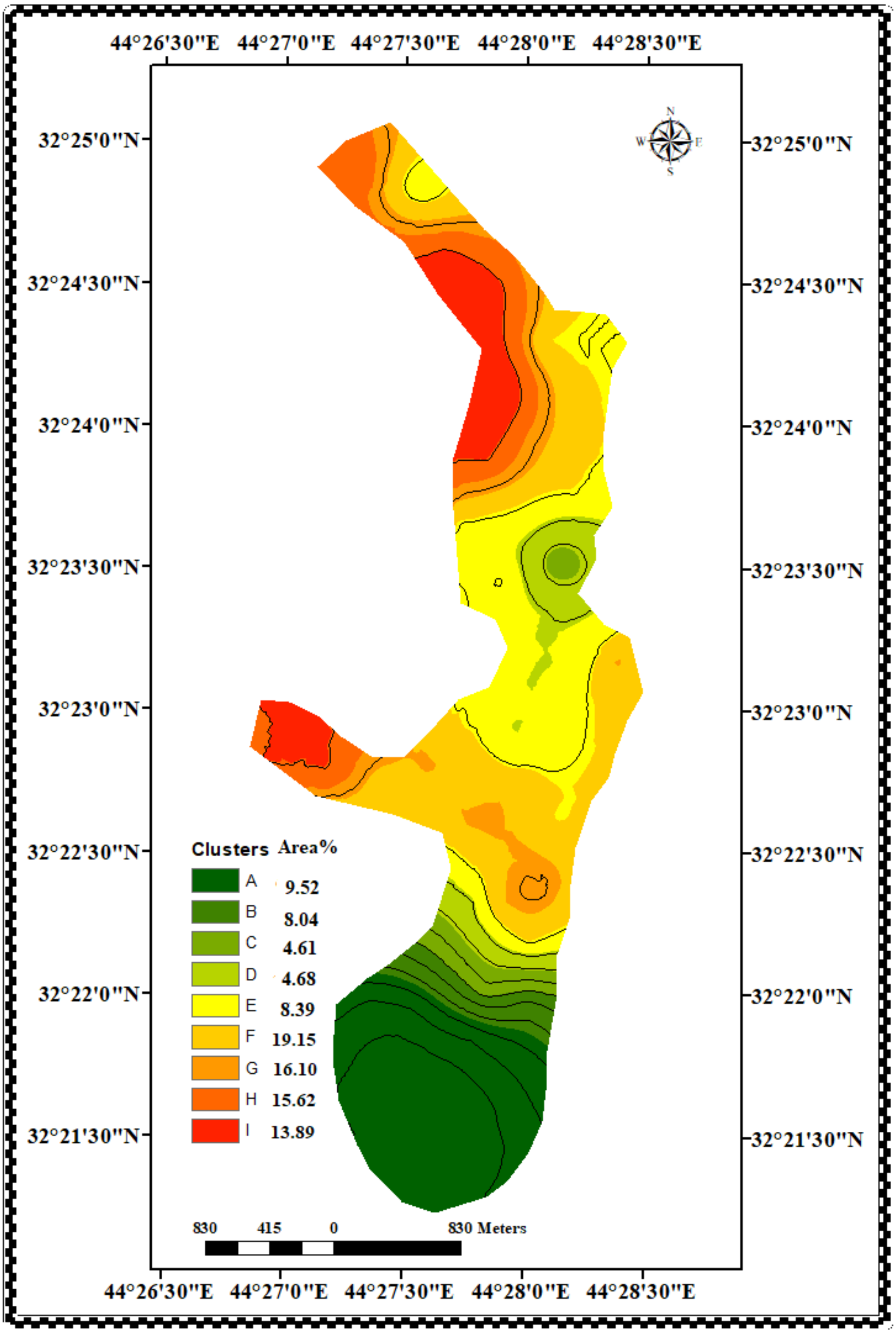

Figure 8. Soil map based on hierarchical cluster analysis for A_1 horizons.

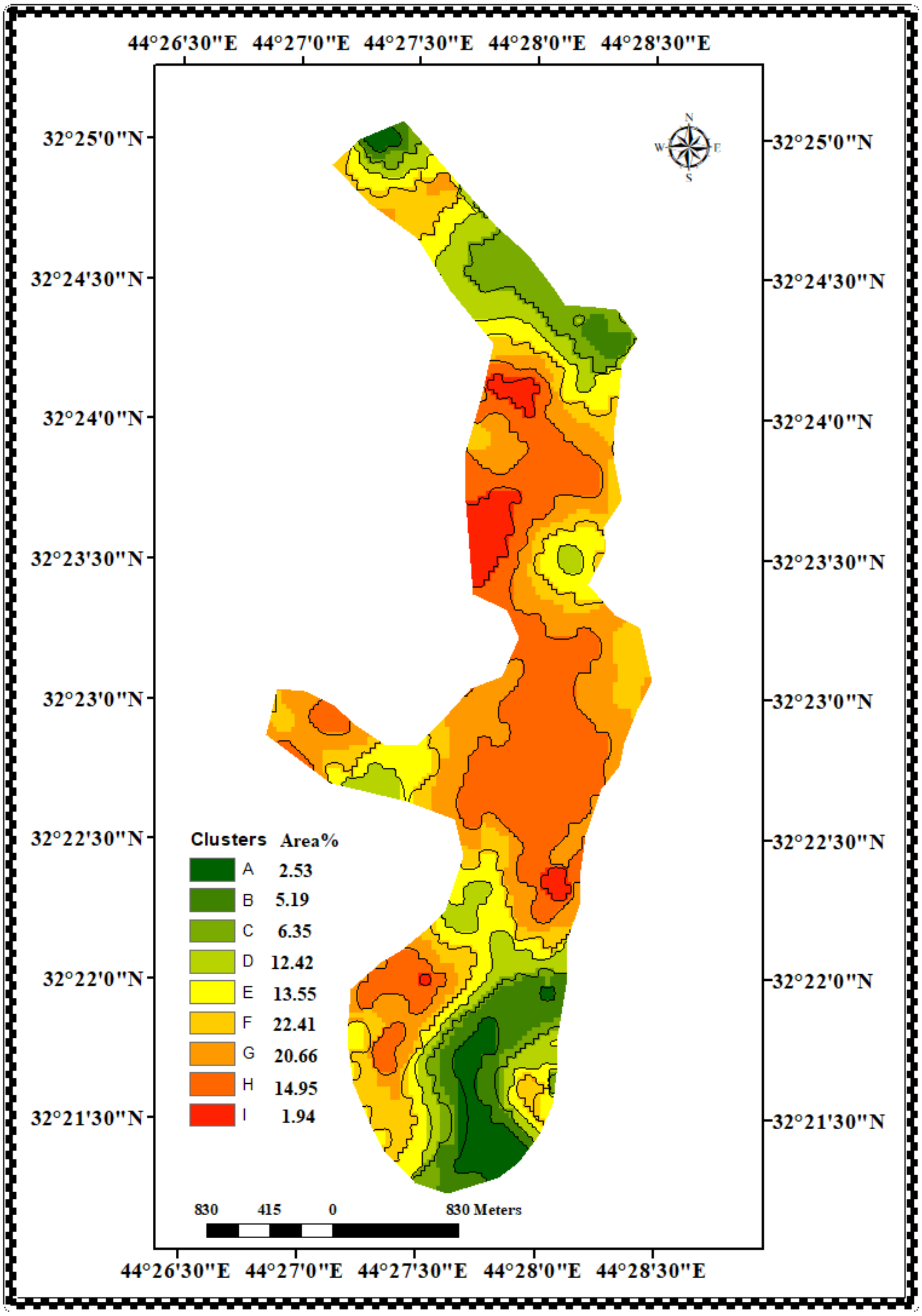

Figure 9. Soil map based on hierarchical cluster analysis for C_1 horizons.

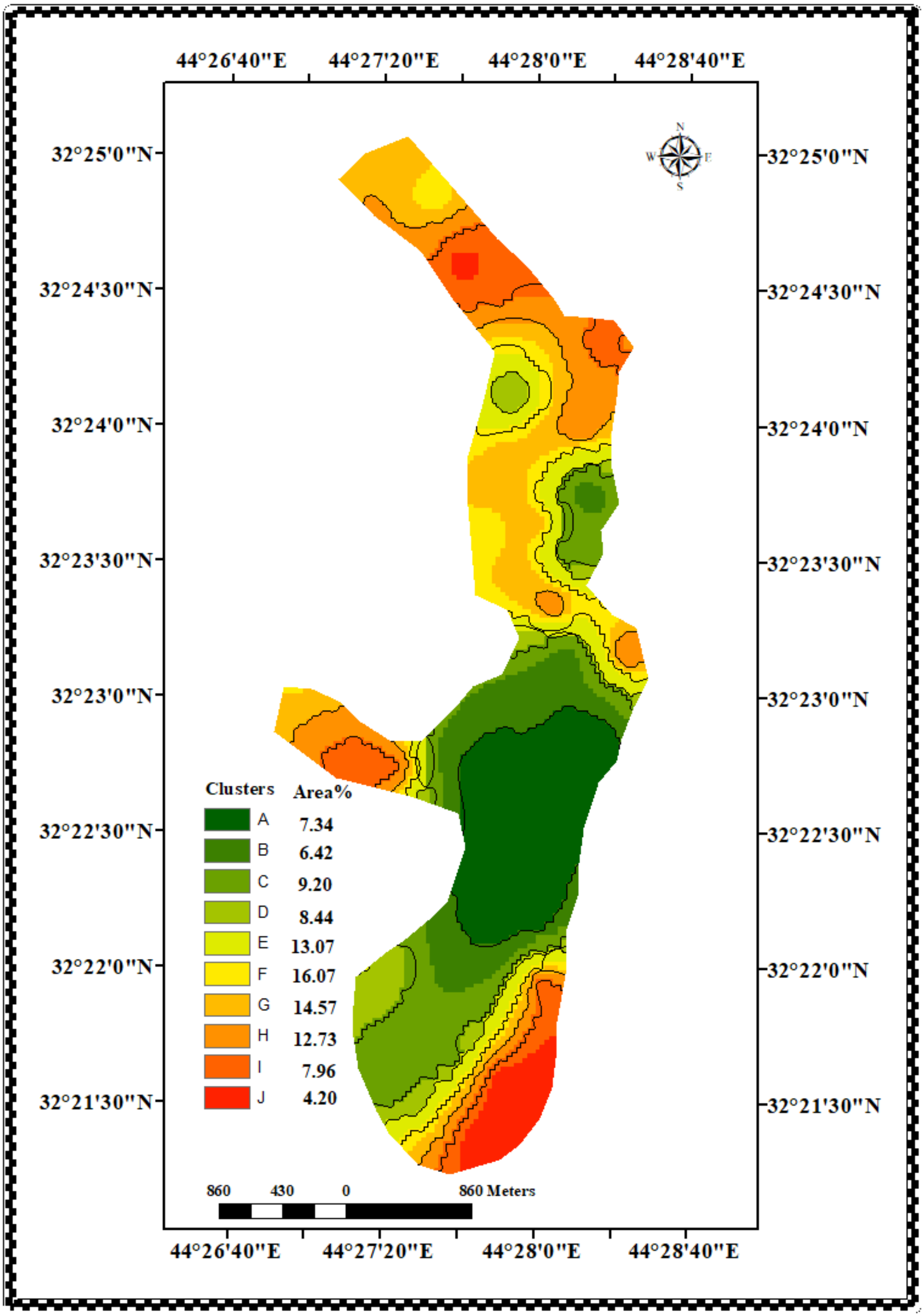

Figure 10. Soil map based on hierarchical cluster analysis for C_2 horizons.

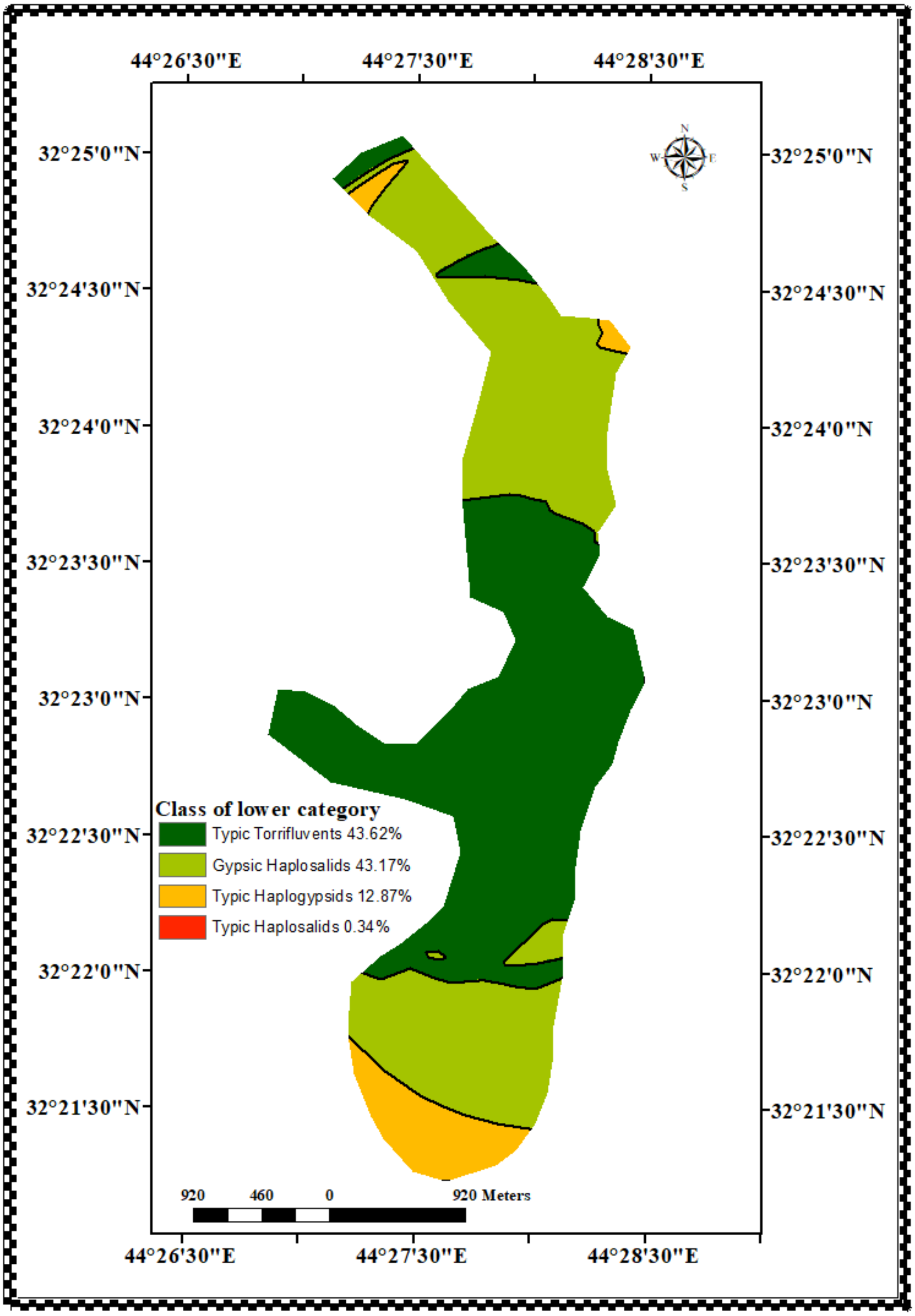

Figure 11. Soil map based on hierarchical cluster analysis for A_1, C_1, and C_2 horizons.

Cluster F dominated at the center of the study area and appeared as a belt in the north - western parts of the study area. Moreover cluster F covered 19.15% of the total study area. Clusters G and H were mainly distributed as narrow belts in the north - western corner of the study area. Further these clusters covered 16.10%-15.62% of the total study area respectively. This type of distribution may be due to the effect of pedogenic processes and to some extent to geomorphic processes, and landscape position (14). The soils which were fell into a given cluster are alluvial in nature, originating from different soils, rocks, unconsolidated sediments deposited by the Tigris and the Euphrates rivers and their tributaries. The Tigris and the Euphrates also carry large quantities of salts. These, too, are spread on the land by sometimes excessive irrigation and flooding (29). The soils based on C_1 soil horizon characters were also classified by hierarchical cluster analysis and compiled into a soil map as shown in Fig. (9). Clusters D and E were relatively scattered all over the area, but their distribution patterns decreased at the center of the study area. In general, these two clusters covered 12.42%-13.55% of the total study area respectively. Cluster F ran across the study area with randomly distribution covered 22.41% of the total study area. While cluster H dominated at the center of the study area and gradually decreased toward north - western corner of the study area. In general, cluster H covered 14.95% of the total study area. The spatial distribution map of these soils reflect the effect of physiographic, geological, and young calcerous alluvium parent materials (18). As a comparison, the soils based on C_2 soil horizon characters were classified by hierarchical cluster analysis and compiled into a soil map as shown in Fig. (10). Cluster A dominated at the center of the study area due to the deposition of the Tigris and the Euphrates river sediments in the direction normal to its flow path (2).

The area of cluster A covered 7.34% of the total study area while cluster F covered 16.07%. appeared as narrow belts in the north and south - eastern parts of the study area. This type of distribution may be due to the effect of climatic conditions, soil texture, landscape positions, groundwater level, quality of irrigation water, and human activities (9). Cluster H was relatively scattered all over the area, covered 12.73% of the total study area, reflecting the effect of the dominant local conditions, mainly, climatic and type of parent materials (15), (17). The soil grouping technique proposed here produced a consistent classification without being affected by an initial setting and led to similar results using different strategies (Fig. 11). We conclude that the proposed method satisfied the practical requirement of the soil classification and mapping and enabled to handle a larger number of soils. The use of cluster analysis, and perhaps other numerical methodologies, can be a useful way to array technologies and methods to identify and quantify soil individuals relationships. Patterns of homogeneity and combinations of distinguishing soil attributes can be more objectively identified through mathematical analyses. We envision cluster analysis as one of many statistical methods that may be used in future soil survey activities. Statistical analyses, combined with careful field observation and evaluation by trained and experienced soil scientists, can produce a new generation of more quantitative soil surveys---

REFERENCES

1.Ahmad, F., M.A.M. Said and L. Najah. 2012. Effect of leaching and gypsum content on properties of gypseous soil. International Journal of Scientific and Research Publications. 2(9): 1-5.

2.Ali, Z. R. and A.S. Muhaimeed. 2016. The study of temporal changes on land cover/land use prevailing in Baghdad governorate using RS and GIS. The Iraqi Journal of Agricultural Sciences. 47(3): 846-855.

3.Allbed, A., and L. Kumar. 2013. Soil salinity mapping and monitoring in arid and semi-arid regions using remote sensing technology: A Review. Advances in Remote Sensing. 2: 373-385.

4.Barazanji, A.F. 1973. Gypsiferous Soils of Iraq. Unpublished Ph.D. Thesis, Ghant, Belgium, pp: 123.

5.Bennett, S.J. and E.G. Barrett-lennard. 2008. Saltland Capability Assessment: Targeting Plants to Landscapes to Increase Profitability, in M. Unkovich, 14^{th} ed., Agronomy

Conference 2008, Sep 21 2008. Adelaide, South Australia: The Regional Institute Ltd. pp: 12.

6.Carré, F. and M. Jacobson. 2009. Numerical classification of soil profile data using distance metrics. Geoderma. 148: 336-345.

7.de Amorim, R.C. 2015. Feature relevance in Ward's hierarchical clustering using the L*p* Norm. Journal of Classification. 32(1): 46-62.

8.Everitt, B.S., S. Landau and M. Leese. 2001. Cluster Analysis. 4th Ed., Oxford University Press, Inc., New York, Arnold, London, pp: 1-2.

9.Ghaderi, R., H. Hamzehzarghani and A. Karegar. 2017. Numerical taxonomy helps identification of Merliniidae and Telotylenchidae (Nematoda: Tylenchoidea) from Iran. The Journal of Nematology. 49(2): 207-222.

10.Gordon, A.D. 1999. Classification, Chapman and Hall/CRC. 2nd ed., pp: 6-10.

11.Legendre, P. and L. Legendre. 2012. Numerical Ecology. 3rd. English Ed., Amsterdam, Elsevier, pp: 12.

12.Martín, M. Á., Y.A. Pachepsky, C. García-Gutiérrez and M. Reyes. 2018. On soil textural classifications and soil-texture-based estimations. Solid Earth. 9: 159-165.

13.Minasny, B. and A.B. McBratney. 2007. Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes. Geoderma. 142: 285-293.

14.Muhaimeed, A.S., A. Ibrahim and R. K. Abdulateef. 2017. Using of remote sensing for monitoring geomorphological teompral changes for Tigris river in Baghdad city. The Iraqi Journal of Agricultural Sciences. 48(1): 215-221.

15.Muhaimeed, A.S. and A.K .Al-Husseiny. 2012. Genesis and development of Argillic horizon in some Northern Iraq soils. The Iraqi Journal of Agricultural Sciences. 43(4): 96-105.

16.Muhameed, A.S. and A.M. Saleh. 2012. Classification of some Iraqi soils using one of the methods proposed in the agglomerative hierarchical clustering. The Iraqi Journal of Agricultural Sciences. 43(4): 1-14.

17.Muhaimeed, A.S. and N.S. Salih. 2009. The effect of local factors on some morphological characteristics at Al-Hamdanya region, Northern of Iraq. Iraq Journal of Agriculture. 19(7): 101-109

18.Muhaimeed, A.S., S.R. al-Jeboory, K.A. Saliem, R. Burt and J.V. Chiaretti. 2013. Genesis and classification of selected soils in an arid region of central Iraq. A peer-reviewed contribution published in Soil Horizons. Soil Science Society of America. 54(6): 1-13.

19.Murtagh, F. and P. Legendre. 2014. Ward's hierarchical clustering method: clustering criterion and agglomerative algorithm. Journal of Classification. 31(3): 274-295.

20.Oliver, M.A. and R. Webster. 1987a. The elucidation of soil pattern in the Wyre forest of the West Midlands, England. I. Multivariate distribution. Journal of Soil Science. 38(2): 279-291.

21.Oliver, M.A. and R. Webster. 1987b. The elucidation of soil pattern in the Wyre forest of the West Midlands, England. II. Spatial distribution. Journal of Soil Science. 38(2): 293-307.

22.Peel, M.C., B.L. Finlayson and T.A. McMahon. 2007. Updated world map of the Köppen–Geiger climate classification. Hydrology and Earth System Sciences.11: 1633-1644.

23.Poshtmasari, H.K., Z.T. Sarvestani, B. Kamkar, S. Shataei and S. Sadeghi. 2012. Comparison of interpolation methods for estimating pH and EC in agricultural fields of Golestan province (North of Iran). International Journal of Agriculture and Crop Sciences. 4(4): 157-167.

24.Rossiter, D.G., R. Zeng and G.-L. Zhang. 2017. Accounting for taxonomic distance in accuracy assessment of soil class predictions. Geoderma. 292:118-127.

25.Smirnova, M.A. and M.I. Gerasimova. 2017. Orders in the soil classification system of Russia: taxonomic distance as a measure of their adequate identification. Eurasian Soil Science. 50(3): 263-275.

26.Soil Survey Staff. 1999. Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys. 2nd ed., USDA-NRCS Agric. Handbook No: 436. US Government Printing Office, Washington, DC, USA, pp: 329.

27.Soil Survey Staff. 2014. Keys to Soil Taxonomy, 12th ed., USDA-NRCS, pp: 131-133.

28.Templ, M., P. Filzmoser and C. Reimann. 2008. Cluster analysis applied to regional geochemical data: problems and possibilities. Applied Geochemistry. 23(8): 2198-2213.
29.Wu, W., A.S. Mhaimeed, W.M. Al-Shafie, F. Ziadat, B. Dhehibi, V. Nangia and E. De Pauw. 2014. Mapping soil salinity changes using remote sensing in Central Iraq. Geoderma Regional. 2-3: 21-31.
30.Young, F. J. and R. D. Hammer. 2000. Defining geographic soil bodies by landscape position, soil taxonomy, and cluster analysis. Soil Science Society of America Journal. 64: 989-998.