J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008

MINIMUM DATA SET APPROACH USING PRINCIPAL COMPONENTS ANALYSIS FOR LAND PRODUCTIVITY ASSESSMENT AT WADI EL-MOLLAK AREA, EAST OF DELTA, EGYPT

MOHAMED E. A. KHALIFA AND ESSAM H. MOHAMED

Pedology Dept., Water resources and desert land division, Desert Research Center, Egypt.

ABSTRACT

Land productivity affected by many permanent or temporal soil advantages and limitations while only few number of soil properties are most variable and mainly affect this productivity. This study aimed at identifying the minimum number of informative soil data required for soil characterizing and potentiality assessment using statistical principal components analysis. Studied area located East of Delta over 11000 feddans of gradually reclaimed soils at Wadi El-Mollak area. An appropriate Land sat scene representing the area was unsupervise classified into 5 spectral classes in which 48 soil profile were distributed. Soil and water samples were collected for laboratory analysis. Fifteen principal components and their proportions were calculated for whole soil data with their related eigenvectors. Results indicated that the first six PCs are the most variance illustrative by 88.925% in which soil salinity, lime content, clay fraction, gravels, profile depth are the highest variable parameters in the studied area. Studied area was classified as *Typic Torrifluevents* for deflated fluvial old deltaic plain soils in which couple of mapping unit were detected and as *Typic Torripsamments* for desert hydrographic sandy basins soils in which three soil mapping unit were distinguished. Results indicated that Good land productivity covered 12.2 % of the area while Fair productivity occupy 51.1 % of the whole area, and Poor productivity was over 34.7 % of the studied area. Limitations were concerning high water table depth, irrigation water salinity, high lime content, low clay

content, high salinity, low organic matter, and low content of NPK. Wheat, Barley, Faba bean, Alfalfa, Grape, Citrus (Orange), Mango, Date palm and Olive are the optimum alternatives in the area.

Keywords: Land evaluation – Land productivity – Principal component analysis – Minimum data set

INTRODUCTION

Agriculture in Egypt represents a milestone in the national economy due to its special historical background, where the cultivated area of Egypt is considered as one of the oldest agriculture areas in the world (Abu Zeid, 1990). Changes introduced in any agricultural national equilibrium result in a number of other changes and precautions ought to be considered to prevent land deterioration. On another hand the population density is one of the highest in the world amounting to almost 1,700 inhabitants per km^2 (Alis, 1991). The major challenge Egypt is facing at the present time is the need for better management of the natural resources to meet the demand of the nation growing at a rate of 2.2% annually (Bishay, 1996).

East of Nile Delta desert remained for a long period not communicated which instigated several enemies before 1973. Both of the military and agricultural strategies were integrated to develop that region which became one of the most promising areas for agrarian production in Egypt. Lots of efforts through many development projects were undertaken such as El-Husienia, El-Salhia, south of Port Saied, and west of Suez canal projects, where the total acreage of reclaimed soils became 0.5 million feddans at East of Delta (Shokry, 1996). Wadi El-Mollak area was reclaimed partially since seventieth, thus a significant difference of land productivity could be easily detected between old and newly reclaimed soils.

Land productivity reflects the quality of soil properties and describes the influence of limiting factors on soil potentiality (Vink, 1975). Lots of philosophies dealt with land productivity in which many soil factors were incorporated and many calculations have to be processed. Minimum data set (MDS) concept describes the minimum information required during processing to produce precise results

J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008

(Matson, 2006). MDS approach can offers an appropriate solution of the frequent question associated with each land evaluation concerning which soil properties is more affecting land productivity?

Principal component analysis (PCA) is one of the most important multivariate statistical analyses which used in reducing and simplifying the attributes data and describe the relationship among several variables (Johnson and Wichern, 1982). This technique employs to select differentiated soil properties which affecting soil variability by reducing the number of individuals without losing important information and arranging them along one or more axes to represent the directions with maximum variability (Webster, 1977).

The current study aimed at creating small number of new variables called principal components using the original soil data and looking among them for the effective soil attributes associated with high variation which used in soil mapping and land productivity assessment in the studied area.

STUDY AREA

1- Location

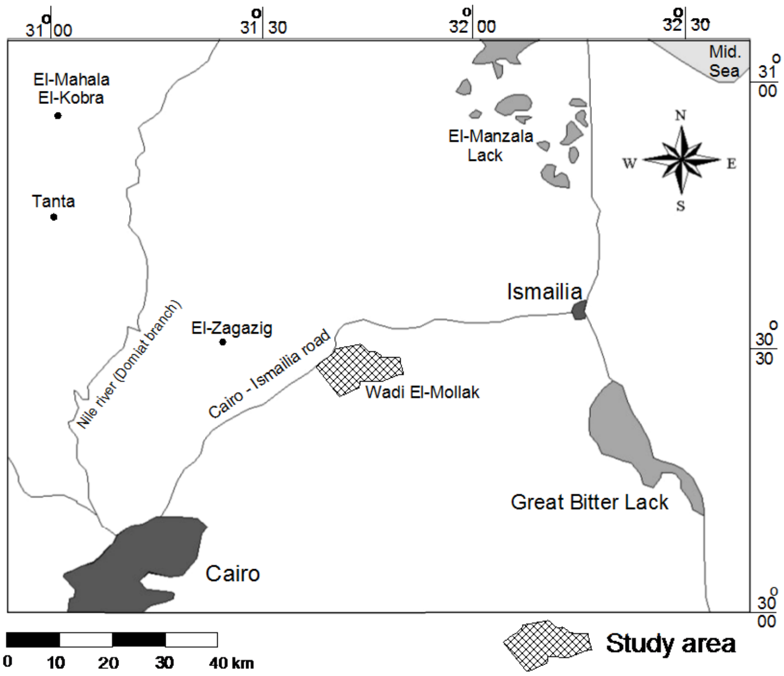
East of Nile delta desert extents towards east direction from Domiata branch to Suez Canal, and towards north direction between Cairo-Suez road to El-Mnzala lake. Studied area comprises part of the newly reclaimed soils in Wadi El-Mollak at East of delta and located between latitudes 30^{o} 25' 30" and 30^{o} 31' 15" and longitudes 31^{o} 40' 30" and 31^{o} 51' 15" and covers about 11000 feds (Map 1).

2- Geology and Geomorphology

According to Geological Survey and Mining Authority (1981), the dominant gravels and sands in the studied area were derived from two sources; the first was the Nile River during the early evolution in past Pliocene time, while the second was the desert hydrographic basins which existed during Miocene time and possibly Quaternary.

East of Delta could be classified into two sectors; the complex southern sector which consists of three rows of calcareous Oligocene and Miocene hills where elevations decreased towards north from 200 to 70 m A.S.L; the plainly northern sector with that great sandy gravely plain which formed over the deflated fluvial old deltaic plain, (Said, 1962). Studied area is belonging to the deflated fluvial old deltaic plain which extends from east of Nile Delta towards Suez

Canal. The plain is characterized by moderately low relief among general flat surface with general northward gentle slopes. Wadi El-Mollak area is one of the developed terraces in the old deltaic plain which formed during the gradual lowering of the Mediterranean Sea. Elevations of the studied area ranged between 35 – 60 m A.S.L.



Map (1) Location of the study area.

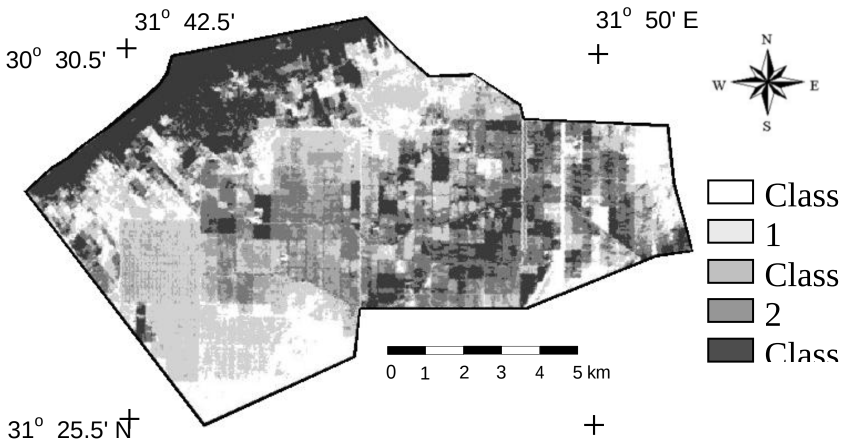
Generally, the plain heights are affected be deposition process of both sands and gravels which was controlled by the changing of Med. Sea level (Abu El-Izz, 1971).

MATERIALS AND METHODS

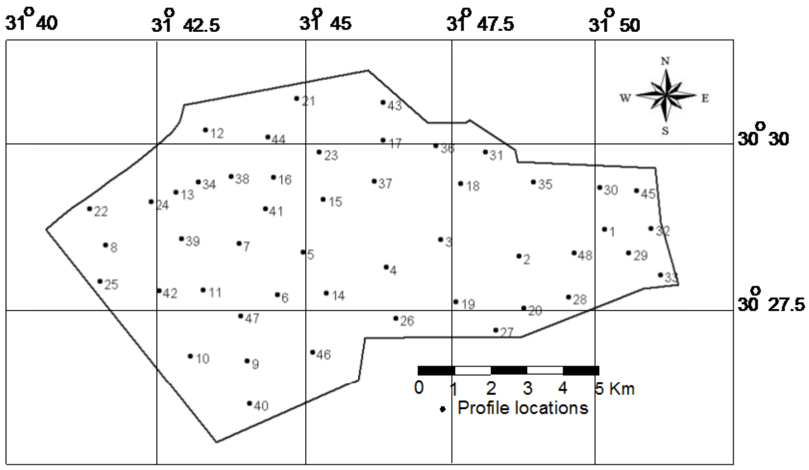
1- Soil survey:

An appropriate satellite image scene (Landsat 7 - ETM+) acquired in 2006 representing the studied area was classified using ERDAS software (Erdas Imagine, 2001). Non significant differences were found by increasing spectral classes, thus the studied area surface was characterized using isodata unsupervised classification by only five classes as seen in map 2.

Studied area was surveyed according to detected unsupervised spectral classes using forty eight representing soil profiles, which were sampled. Ten water samples were collected from either Ismeilia or El-Molak canals branches at different sites, in addition to five wells water samples. The observations locations were georeferenced to the UTM coordinate system (ESRI, 2006) for further processing and presented as shown in map 3.



Map (2) Isodata unsupervised classification of the study area.



Map (3) Remotely-sensed design of soil sampling in the study area.

2- Laboratory analysis:

Collected soil samples were analyzed to determine (1) physical properties (Page et al., 1982); saturation percent, soil texture, gravel, available water and hydraulic conductivity, (2) chemical properties (Page et al., 1982); electrical conductivity of past extracts (EC) in dS/m, soluble cations and anions in $meql^{-1}$, soil reaction (pH), total calcium carbonate and gypsum content (Jackson, 1973), (3) fertility properties; organic carbon (Jackson, 1973), total nitrogen (Black, 1983), available phosphorus and potassium (Soltanpour, 1985), available micro elements (Lindsay and Norvel, 1978). Water samples were analyzed to determine water salinity, soluble cations & anions, pH, boron and nitrate (Page et al., 1982).

3- Statistical principal component analysis

Statistically, PCA is used as an unbiased method to explain the variance structure through a few linear combinations of the original variables, each combination called principal component (PCs) (Afifi and Clark, 1984 & Richard and Wichern, 1975). Geometrically, these linear combinations represent the selection of a new coordinate system obtained by rotating the original system which uses studied variables X1, X2, ..., Xi as coordinate axes. Each linear combination, describes one of (PCs), can be written as:

PCi = Ai1 X1 + Ai2 X2 + Ai3 X3 + ... + Ain Xn

Where PCi: principal component, Ain: chosen coefficient called "eigenvectors" or characteristic vectors of a matrix which used in creating the principal components (PCs) from the original variables Xn. (Ovalles and Collins, 1988).

The samples (N) and the variables (P) form a matrix A= [N*P]. The characteristic roots of this matrix called "eigenvalues" which used in calculating the proportion of each component, and then the cumulative proportion.

$$\text{Proportion of } PC_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \ldots + \lambda_n}$$

where n : the number of components.

The number of selected principal components for studying was determined by using a rule of thumb (Afifi and Clark, 1984), where selected PCs are those that explain at least 100 / P percent of the total variance for each of them, where P is the number of original variables. Selected PC, which has a large loading in the component, must has an eigenvector value larger than a calculated threshold called "selection criterion" (SC) where:

$$SC = 0.5 / (PC\ eigenvalue)^{½}$$

Finally, variables located within selected components represent the most variation in the population. SYSTAT package was used to calculate the eigenvalue and eigenvectors of each component.

4- Land Evaluation

According to PLES ARID (Khalifa 2004) based on Sys (1991-I&II and 1993) and El-Fayoumy (1989), Soil data were rated to produce land productivity indices of the studied area. Evaluated parameters include: soil physical properties, chemical properties, topography, fertility elements, water quality and climate characteristics. Land Suitability for different crops were also investigated.

RESULTS AND DISCUSION

1- Statistical analysis

Table (1) represents the eigenvalues or the characteristic roots of samples-variables matrix with their proportions corresponding to fifteen calculated principal components. The proportion reflects the variance introduced through information conveyed by each principal component (PC). Thus, PCs were ordered decreasingly with respect to the variance. First PC is the most informative component, whereas the last is the lowest informative one. The selected PCs to explain the variance are those in which proportions have to exceed 100 / number of variables (100 / 26 ≈ 4) percent of the total variance. Therefore, the selected PCs which have eigenvalues proportions more than 4 % are the first six PCs which explained 88.925% of the cumulative variance.

J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008

Table (1) Eigenvalue and Proportion of total variance explained by each principal component for weighted average data.

Principal Component no.	Eigenvalue	Proportion (%)	Cumulative proportion (%)
1	7.182	35.526	35.526
2	6.548	20.271	55.797
3	3.524	12.497	68.294
4	2.025	8.101	76.395
5	1.840	7.373	83.768
6	1.249	5.158	88.925
7	0.732	2.926	91.852
8	0.487	1.947	93.799
9	0.429	1.716	95.515
10	0.301	1.204	96.719
11	0.289	1.156	97.875
12	0.236	0.946	98.821
13	0.115	0.572	99.393
14	0.028	0.399	99.792
15	0.008	0.208	100.000

Eigenvectors which reflect the loading (the weight) of each studied variable within a component were calculated as seen in table 2. Selection criterions for selected principal component were also calculated (table 2) and compared with eigenvectors inside each component to identify larger eigenvectors corresponding to highly weighted soil properties with most affection in creating soil variability.

Table (2) indicated that PC1 which had 33.25% from the total variance contained the higher number of the selected variables than in other selected components. Soil salinity, calcium carbonate, clay content, gravels and profile depth are the most variable parameters due to their appearance four times and also due to their larger loading (eigenvectors) in their components. Selected characteristics are responsible about soil variability in the studied area.

J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008

Table (2) Eigenvectors of the correlation matrix for weighted soil properties of the studied area.

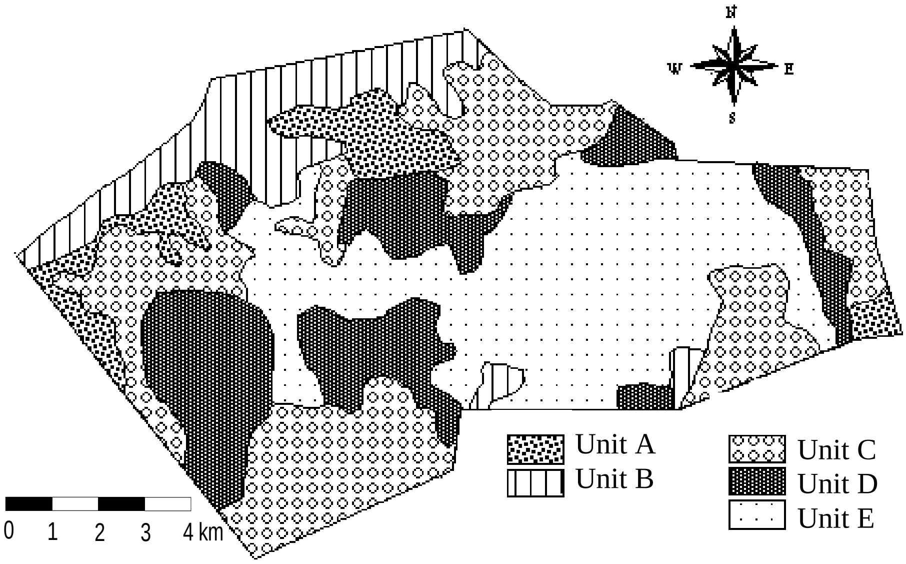
Soil Property	Principal component					
	1	**2**	**3**	**4**	**5**	**6**
pH	0.193	-	0.325	-	-	0.669
EC	0.714	0.619	0.899	-	-	0.541
$CaCO_3$	0.863	0.717	-	0.714	0.727	-
Gypsum	0.684	0.722	-	0.613	-	-
Sand	0.355	0.772	-	-	-	-
Silt	0.417	-	-	0.579	-	-
Clay	0.633	0.700	0.843	-	-	0.652
A.W.	0.404	0.763	-	-	-	-
Kh	0.735	0.556	-	-	-	-
Gravels	0.635	0.656	0.754	0.475	-	-
S.P	0.256	0.412	0.549	-	-	-
O.M.	-	-	0.369	0.630	-	-
N	-	-	0.505	0.597	-	-
P	-	-	0.355	-	0.455	-
K	0.458	0.540	-	-	-	-
Depth	0.606	0.743	0.669	-	0.394	-
S.C.*	0.173	0.212	0.308	0.315	0.368	0.440

* S.C. = Selection criterion.

2- Soil mapping units

Table (3) summarized some selected soil properties of represented profiles in the studied area. Map (4) represents different soil mapping units of the studied area at Wadi El-Mollak area, which classified regarding to soil profile depth, salinity, calcium carbonate content, and texture which have been identified during PCA. Generally, Soils of the studied area were derived and formed mainly from (1) deflated fluvial old deltaic plain for finer soils and from (2) desert hydrographic basins for wind blown sands and gravels sediments. Five soil mapping units were identified: unit A over 8.6%, unit B over 12.2, unit C over 24.5%, unit D over 18.0% and unit E over 34.7% of the total area (map 4).

J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008



Map (4) Soil mapping units of the studied area at Wadi El-Mollak

(1) Soils of deflated fluvial old deltaic plain
A) Deep sandy clay loam moderately saline slightly calcareous
B) Moderately deep clay loam slightly saline non calcareous
(2) Soils of desert hydrographic sandy basins:
C) Deep sandy clay loam gravely moderately saline moderately calcareous
D) Deep sandy clay loam saline moderately calcareous
E) Very deep sandy loam gravely moderately saline calcareous

Laboratory analyses and field investigations indicated that soils of the studied area following Entisols order and sub classified as *Typic Torripsamments* and *Typic Torrifluevents*.

3- Land Productivity

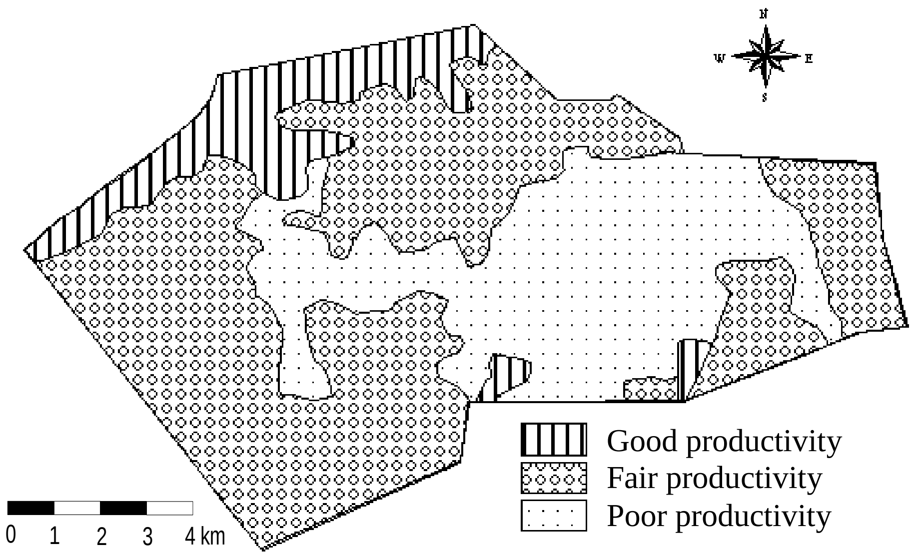
Three land productivity classes were identified in the studied area as seen in table (4) and map (5). Good land productivity occupies 12.2 % from the studied area for soil unit B, Fair land productivity cover an area of 51.1 % of the studied location which representing soil units A, C & D, while Poor land productivity was found over 34.7% of the total area which belong to soil unit E. Limitations in the area

J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008

may be concluded as high water table depth, saline irrigation water, high lime content, low clay fraction, high salinity, low organic materials, and low content of nitrogen, phosphorus and potassium.

Table (4) Land productivity classes in the studied area.

Soil Mapping Unit	Phys. index	Chem. index	Topo. index	Fert. index	Water index	**Final index**	Land Productivity Class
A	54.4	74.2	84.5	42.5	85.4	**68.20**	Fair
B	60.2	82.0	88.2	45.8	87.9	**72.82**	Good
C	45.2	74.2	75.6	28.2	72.8	**59.20**	Fair
D	62.0	44.8	72.0	27.4	75.7	**56.38**	Fair
E	45.2	35.6	62.7	18.8	80.0	**48.46**	Poor



Map (5) Land productivity classes of the studied area

An associated land suitability study was performed for optimum land use in the studied area. Results recommend some highly suitable alternative crops at Wadi El-Molak area, these are Wheat, Barley, Faba bean, Alfalfa, Grape, Citrus (Orange), Mango, Date palm and Olive.

REFERENCES

Abu El-Izz, M. S., (1971). Land Form Of Egypt, John Wiley And Sons, New York.

Abu Zeid, M. A. (1990). "Strategies Of Sustainable Development In Egypt," Summary Report, Dr. Adli Bishay,

Afifi, A. A. and Clark, V. (1984). Computer-aided multivariate analysis. Lifetime Learning Pub. Belmont, CA.

Alis, S. (1991). "Agricultural Land information System Of Egypt. " Agricultural research center . Giza, Egypt.

Bishay, A. (1996). " Desert Development For Our Future Generations". The Fifth International Desert Development Conference, Lubbock, Texas"12-17 Augast, 1996."

Black, C. A. (1983). Methods of Soil Analysis, part I and II Soil Sci.An.Inc. Public. Madison , Wise., USA.

ElFayoumy, M.E., (1989). New approach of land evaluation of some Egyptian regions. Ph.D. Thesis, Fac. of Agric., Univ. of Alexandria.

ERDAS IMAGINE 8.6. (2001). Geographic imagine package. ERDAS Inc. USA.

ESRI. (2006). Arc-GIS user manual. Version 9.2. Redlands. California.

Geological Survey and Mining Authority, (1981). Geological Map Of Egypt., Scale 1:2.000.000, Cairo, Egypt.

Jackson, M.L., (1973). Soil chemical analysis. Advanced course Ed.2. A Manual of methods useful for instruction and research in soil chemistry, physical chemistry of soil, soil fertility and soil genesis. Revised from Original Edition (1955).

Johnson, A. R. and Wichern. (1982). Applied multivariate statistical analysis. Prentic Hall, Inc., Englwood Cliffs, NJ, USA.

Khalifa, M. E. A., (2004). An Integrated Parametric Evaluation Model for Land Resource Management: A Case Study for El-Bostan Extension Sector, West Nubaria , Egypt. Ph.D Thesis, Soil & Water Dept., Faculty of agriculture, University of Alexandria, Egypt.

Lindsay, W.L. and Norvell, W.A. (1978). Development of DTPA Soil Test for Zn, Fe, Mn and Cu. Soil Sci. Soe. Am. J., Vol.42:421-428.

J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008

Matson, J. (2006). Minimum Data Set – the digital solution for complex data bases, Moracon Pub., NY, USA

Ovalles, F. A. and Colilins, M. E. (1988). Variability of Northwest Florida soils by principal component analysis, Soil Sci. Soc Am. J., 52:1430-1435.

Page, A. L., Miller, R. H. and Keeny, D. R. (1982). Methods of Soil Analysis, Part 2- Chemical and Microbiological Properties. Agronomy Monograph No. 9. ASA, SSSA, Madison, WI.

Richard, A. J. and Wichern, D. W. (1975). Applied multivariate statistical analysis Uni. Wisconsin – Madison, New Jersey, USA

Said, R., (1962). "The Geology Of Egypt." Elsevier Publishing Co. Amsterdam, New York.

Shokry, N., (1996). Egypt's Agriculture Investment Document during Mobarak's Age, published by Eldawalia House, handbook no. 11337. (in Arabic)

Soltanpour, P. N. (1985). Use of Ammonium Bicarbonate DTPA Soil Test to Evaluate Elemental Availability and Toxicity. Common. Soil Sci., Plant Anal., Vol. 163:323-338.

Sys, C.; Van Ramst, E.; and Debaveye, J. (1991). Land evaluation. Part I – Principle in land evaluation and crop production calculation. Agricultural publication - no.7, General administration for development cooperation, Brussels, Belgium.

Sys, C., Van Ramst, E., Debaveye, J., and Beernaert, F., (1993). Land evaluation. Part III – Crop requirements. Agricultural publication - no. 7, central administration for development cooperation, Brussels, Belgium

Vink, A. P. A. (1975). Land use in advancing agriculture springer, Berlin.

Webster, R., (1977). Quantitative and Numerical method in soil classification and survey. Clarendon Press, Oxford.

J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008

الملخص العربي

استخدام تحليل المكونات الرئيسية لتحديد الحد الأدنى من البيانات لتقييم قدرة التربة الإنتاجية بمنطقة وادي الملاك – شرق الدلتا – مصر.

محمد عزت عبد الهادى خليفة – عصام حسين محمد

قسم البيدولوجى – شعبة مصادر المياة والأراضى الصحراوية – مركز بحوث الصحراء – مصر.

تعد القدرة الإنتاجية للأراضى محصلة لمجموعة مقومات ومعوقات إستغلال التربة الثابتة أو المتغيرة، وتقييم الأراضى المروية يهدف فى النهاية إلى زيادة الإنتاج الزراعى. ولهذا فإن تقييم قدرة التربة الإنتاجية يستلزم حصر العديد من متغيراتها للوقوف على المتدنى منها، إلا أنه سيظل عددا محدودا فقط من صفات التربة هى الأكثر تباينا وتأثيرا على تحديد قدرة التربة الإنتاجية. وتهدف هذه الدراسة إلى تحديد الحد الأدنى من بيانات التربة المسئول عن تعريف أنواع وحدات التربة الاساسية ودرجات قدرتها الإنتاجية وذلك بإستخدام طريقة تحليل المكونات الأساسية الإحصائية. منطقة الدراسة أختيرت بإقليم شرق الدلتا على إمتداد مساحة 11 ألف فدان تقريبا بمنطقة وادى الملاك لتمثل الأراضى الحديثة والتى تم إستصلاحها على مراحل متتالية. وقسمت صورة القمر الأمريكى لاندسات الممثلة للمنطقة طيفيا بإستخدام التقسيم الغير موجه لعدد 5 أقسام وزع بها عدد 48 قطاع لحصر أراضى المنطقة وتجميع عينات التربة الممثلة، إضافة لتجميع عدد 15 عينة مائية من فروع ترعة الإسماعيلية والملاك وبعض الأبار الجوفية بالمنطقة. وتعرضت كافة العينات للتحليل المعملى، ثم حللت بيانات التربة إحصائيا بإستخدام تحليل المكونات حيث أمكن حساب عدد 15 مكون أساسى PC إضافة لنسب إستحوازها Proportion شاملا لكافة صفات التربة ونسب تأثيرها Eigenvectors. وأظهرت النتائج أن المكونات الأساسية الست الأولى قد فسرت 88.925% من تباين صفات التربة وأن صفات ملوحة التربة ونسبة الجير ونسبة المكون الطينى (قوام التربة) ونسبة الحصى وعمق القطاع الأراضى هى الأكثر ظهورا بها ومن ثم يجب الإعتداد بها فقط لتمييز وتقييم الأنواع الأرضية. قسمت منطقة الدراسة لتتبع رتبة الإنتيسول وصولا لتحت المجموعة *Typic Torrifluevents* لتميز أراضى السهل الدلتاوى القديم و *Typic Torripsamments* مميزا لأراضى أحواض الترسيب الصحراوية، كما أمكن تمييز عدد 5 أنواع أرضية بداخل هذه الوحدات تباينت فى الصفات سابقة التحديد. أوضحت النتائج أن قدرة التربة الإنتاجية كانت جيدة على إمتداد مساحة 12.2 % من منطقة الدراسة، وتدنت القدرة الإنتاجية للمستوى المتوسط لأكثر من نصف المساحة المدروسة على إمتداد 51.1 % منها، بينما شغلت قدرة التربة الفقيرة مساحة 34.7 % من إجمالى المساحة. وإنحصرت معوقات الإستغلال ببعض مساحات منطقة الدراسة فى إرتفاع مستوى الماء الأرضى، إستخدام ماء رى ملحى خصوصا ببعض آبار جنوب المنطقة، إرتفاع نسبة الجير، إنخفاض نسب المكون الطينى، إرتفاع الملوحة نسبيا، التدنى العام لعناصر خصوبة التربة. أوصت الدراسة بضرورة تعميم إستزراع محاصيل القمح، الشعير، الفول، البرسيم، العنب، البرتقال الصيفى، المانجو، نخيل البلح، الزيتون .

J.Agric.&Env.Sci.Alex.Univ.,Egypt Vol.7 (3)2008