EnvironmentAsia
The international journal published by the Thai Society of Higher Education Institutes on Environment

Available online at www.tshe.org/EA
EnvironmentAsia 3(1) (2010) 47-52

Soil Classification Based on their Chemical Composition Using Principal Component Analysis

Kamolchanok Panishkan, Mayuva Areekijseree, Natdhera Sanmanee and Kanokporn Swangjang

Faculty of Science, Silpakorn University, Nakorn Pathom, 73000 Thailand

Abstract

Principal component analysis (PCA), was used to group of 54 soil samples collected from different agricultural locations in the western part of Thailand. Soil chemical compositions were measured by Scanning Electron Microscope (SEM) and Energy Dispersive X-ray Microanalysis (EDX). The basic result indicated decreased amount of O, Si, Al, Fe, C, K, Mg, Ca, Ti and Na. The first three principal components were used and accounted for 44.0%, 19.5% and 15.0% of total variation of the data, respectively. Score plots of first three principal components were used to map with soil textures classified as clay, clay loam and medium loam. The results showed some relationships between chemical contents and soil textures. PCA was shown to be a useful tool for soil texture classification based on their chemical compositions.

Keywords: principal component analysis; agricultural soils; Thailand.

1. Introduction

Principal component analysis (PCA) is a multivariate analysis technique and is usually applied in environmental and agricultural studies (Tariq *et al*., 2005; Boruvka *et al*., 2005; Manzoor *et al*., 2006; Dragovic and Onjia, 2006; Mico *et al*., 2006). In this work, PCA was used to cluster patterns of soil samples collected from different agricultural locations in Thailand based on their chemical compositions. Soil chemical compositions were measured by Scanning Electron Microscope (SEM) and Energy Dispersive X-ray Microanalysis (EDX).

The aims of this study were (1) to determine contents of oxygen (O), magnesium (Mg), aluminum (Al), potassium (K), silicon (Si), calcium (Ca), titanium (Ti), Sodium (Na), Carbon (C) and Iron (Fe) in soil samples (2) to identify their relationships by using the Pearson correlation and (3) to study classification of soil samples according to their chemical characteristics by using principal component analysis (PCA).

2. Materials and methods

2.1. Soil Sampling and Preparations

Fifty four agricultural soils obtained from 3 provinces, Nakorn Pathom, Samut Sakhon and Samut Songkhram were included in this present study. Soil samples were randomly collected 7-10 points at each location at 15 cm-depth and mixed. Then, one-fourth of the mixed soil approximately 3-4 kilograms were brought back and air dried at the laboratory. Then, they were sieved pass through nylon sieve pore size 0.65 um. Chemical parameters such as soil color and pH were measured after Sparks *et al*. (1996). Particle-size analysis was done by hydrometer method and soil texture was classified after United States Department of Agriculture (USDA) (Gee and Bauder, 1986). For USDA classification, sands were <2000-50 m, silts were <50-2 m, and clays were <2 m. The different ratios of these 3 sizes of soil particles indicated different types of soil texture (Fig. 1).

For Scanning Electron Microscope (SEM) and Energy Dispersive X-ray Microanalysis (EDX), each soil sample was weighted (0.2 g) and prepared by Lab Press (International Crystal Laboratories, USA). The samples were flat and smooth specifically for quantitative and qualitative determination. Then, soil samples were mounted on stubs with conductive carbon tape. All soil samples were analyzed for their micro-structure and elemental compositions under SEM (JOEL JSM 6400 LV) operating at 15 kV.

The analysis on the elemental compositions of the soil samples have been attempted using SEM/EDX. EDX has limits of 0.1% (approximate). The percentage of element ranging from 10^5-10^6 atoms is being detected by this method (Jenkin *et al*., 1995).

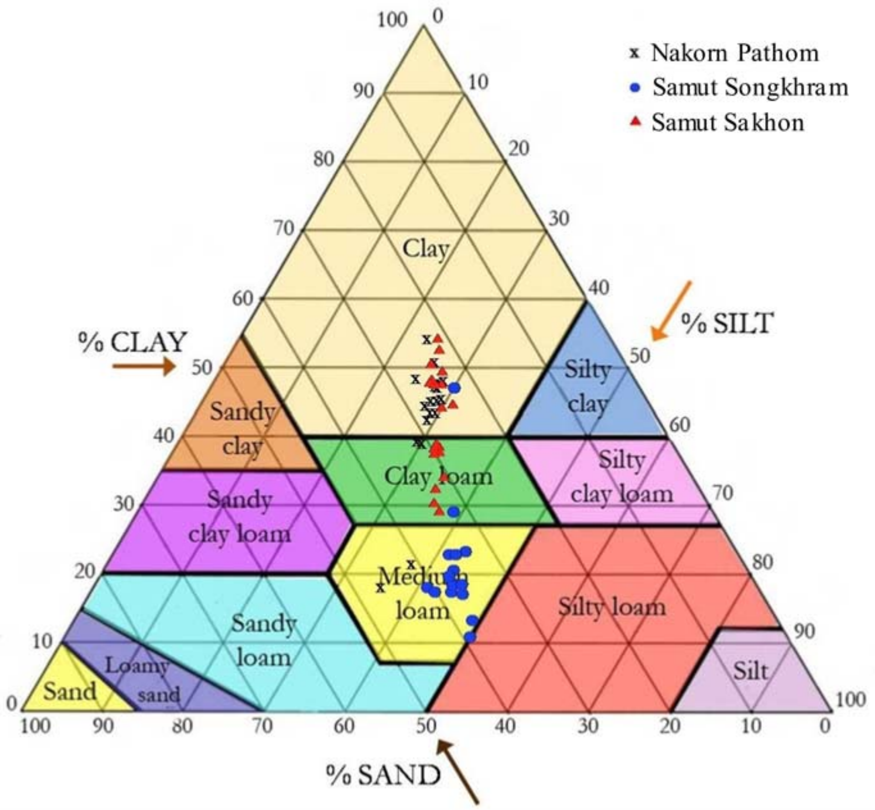


Figure 1. Soil textural triangle. (x) = soil samples obtained from Nakorn Pathom province, (l) = soil samples obtained from Samut Songkhram, and (s) = soil samples obtained from Samut Sakhon.

2.2. Statistical Analysis

The data was analyzed based on 54 soil samples (26 of Clay and 12 of Clay Loam and 16 of Medium Loam) collected from the three study areas. Descriptive statistics, including mean, standard deviation, minimum and maximum percentages of O, Si, Al, Fe, C, K, Mg, Ca, Ti and Na were performed. Simple correlations were analyzed to determine the relationships between different soil parameters. PCA based on correlation matrix was used to examine patterns between soil chemical compositions and soil textures.

2.3. Principal Component Analysis (PCA)

Investigate often measure or make observations on a large number of variables. There are several useful techniques to reduce the dimensionality of data without the loss of much information. Principal component analysis is one such technique. Typically, principal component analysis is used to reduce the dimensionality of a data set, while retaining as much of the original information as possible. This is achieved by transforming the original set of variables into a smaller set of linear combinations called principal components (Jolliffe, 1986; Rencher, 1995). Their values are called the component scores and the coefficients of linear combinations are called loadings. The components are uncorrelated and account for the total variance of the original variables. The first principal component (PC1) has the largest variance and accounts the greatest amount of the total variance. The second principal component (PC2) has the second largest variance and contributes the greatest amount of the residual variance, and so on. The score plot is used for examining the classification of the data. The loading plot is used for investigating the importance of variable to each component.

3. Results and Conclusions

3.1. Soil Characteristics

Fifty four agricultural soils obtained from 3 provinces distributed into 3 types; 26 soils as clay, 12 soils as clay loam, and 16 soils as medium loam (Fig. 1). Their colors were in yellow-red type (data not shown). Most of them had a pH in the range of 6.5-7.5.

3.2. PCA Results

Table 1 showed the basic results of soil chemical compositions indicated the decreased amounts of O,

Table 1. Basic statistics of the distribution of selected soil chemical parameters (%)

	Minimum	Maximum	Mean	Std. Deviation
O	50.390	56.910	54.395	1.595
Mg	0.290	1.170	0.796	0.159
Al	7.550	10.410	8.709	0.723
Si	22.820	31.750	26.276	2.393
K	1.270	2.370	1.732	0.289
Ca	0.000	1.330	0.581	0.252
Fe	2.780	5.080	3.884	0.558
C	0.000	8.350	2.920	2.825
Ti	0.000	0.530	0.368	0.103
Na	0.000	0.380	0.042	0.087

Si, Al, Fe, C, K, Mg, Ca, Ti and Na. The contents of O, Si and Al presented higher levels indicating the great amounts of clay mineral in all soils, whereas Mg, Ca, Ti and Na presented the lowest levels representing some minute amounts of imbedded minerals in the soil textures. The amounts of O, Si and Al ranged from 50.39%-56.91%, 22.82%-31.75% and 7.55-10.41%, respectively.

The Pearson correlation matrix in Table 2 indicated a number of strong associations between pairs of the ten soil parameters. The results showed the strong positive associations between all pairs of Al, Si and Fe (Fig. 2). It implied that alumina and silica sheets exhibiting octahedral and tetrahedron units with oxygen of clay minerals (McBride, 1994) were the main minerals found in all soils. Thus, the decreasing of oxygen might come from replacing atoms of Al, Si, K, and Fe in the clay mineral structures as shown by their significant negative correlations (*r = -0.442, -0.571, -0.339* and *-0.696*, respectively). Another important constituent in soil was organic matters. Carbon found in soils mostly derived from organic matters in which it contained oxygen in its structure as well (Sposito, 1989; McBride, 1994; Stevenson, 1994). As it appeared in soils, the lesser fractions of others occurred. Therefore, the negatively correlations with other elements were shown especially Al, Si, K and Fe (*r = -0.764, -0.883, -0.428* and *-0.670*, respectively).

Table 2. Pearson correlation coefficients between 10 elemental compositions in soil

	O	Mg	Al	Si	K	Ca	Fe	C	Ti
Mg	-0.002								
Al	-0.442*	-0.071							
Si	-0.571*	0.182	0.788*						
K	-0.339*	0.376*	0.361*	0.490*					
Ca	0.053	-0.149	-0.051	-0.018	-0.498*				
Fe	-0.696*	0.258	0.624*	0.809*	0.347*	-0.014			
C	0.183	-0.269*	-0.764*	-0.883*	-0.428*	-0.027	-0.670*		
Ti	0.173	0.584*	-0.156	-0.104	0.071	0.141	0.011	-0.080	
Na	0.135	0.463*	-0.100	0.125	0.312*	-0.301*	0.129	-0.228	0.257

* Significant at $P<0.05$.

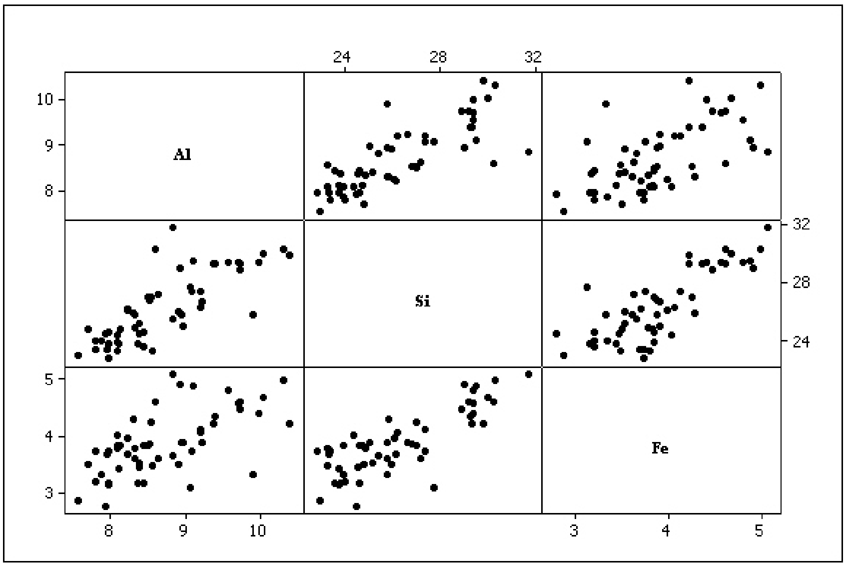


Figure 2. Relationships between Al, Si and Fe

The results of PCA are presented in Table 3. The PCA loadings>4 are shown in bold. The first three principal components were obtained with eigenvalues>1 and accounted for 44.0%, 19.5% and 15.0% of total variation of the data, respectively, cumulatively contributing 78.5% of variance. The first PC showed high loadings of Al, Si and Fe with positive effect and C with negative effect. The second PC was associated with Mg and Ti with both negative effects. The third PC included K and Ca with highly negative loading on Ca content. A 3-D plot of loadings in Fig. 3 and score plots in Fig. 4 showed some relationships among soil contents, clay, clay loam and medium loam. From the first 3 principal components (PC1, PC2 and PC3) mapping with soil textures in Fig. 4, the distributions of clay were found mostly around the center implying similar influence of all 3 principal components to this type of soil. As the amounts of sand and silt increased for clay loam and medium loam (Fig. 1), the distributions of 3 principal components of these 2 types of soil were more scatter. Yet, some trends were shown and could be used to differentiate these two groups from one another. For clay loam in which it contained high amounts of organic matters (high C), it showed lower cluster of PC1 than medium loam in which only some soils of this type that were regularly fertilized with organic fertilizer. The higher PC3 was also found in medium loam corresponding to the higher K as it was widely used for enhancing the sweetness of fruits at Samut Songkhram province (Fig. 1).

In conclusion, the PCA helped to reveal some relationships between chemical contents, soil textures and agricultural practices. It was shown to be a useful tool for soil study and investigate what might have caused based on their chemical compositions. More data would provide more reliable prediction among those links between them.

Table 3. Loadings of the first 5 components for elemental compositions in soil

compositions	PC1	PC2	PC3	PC4	PC5
O	-0.322	-0.203	-0.097	**-0.739**	0.187
Mg	0.137	**-0.655**	-0.048	0.137	**0.486**
Al	**0.415**	0.211	-0.068	-0.294	-0.315
Si	**0.480**	0.079	-0.098	-0.090	0.111
K	0.309	-0.269	**0.455**	-0.059	**-0.452**
Ca	-0.074	0.160	**-0.767**	0.118	-0.170
Fe	**0.439**	0.030	-0.134	0.309	0.313
C	**-0.426**	0.064	0.216	**0.466**	-0.127
Ti	-0.014	**-0.612**	-0.337	0.105	**-0.520**
Eigenvalue	3.960	1.751	1.353	0.87	0.400
Cumulative percentage	0.440	0.635	0.785	0.882	0.926

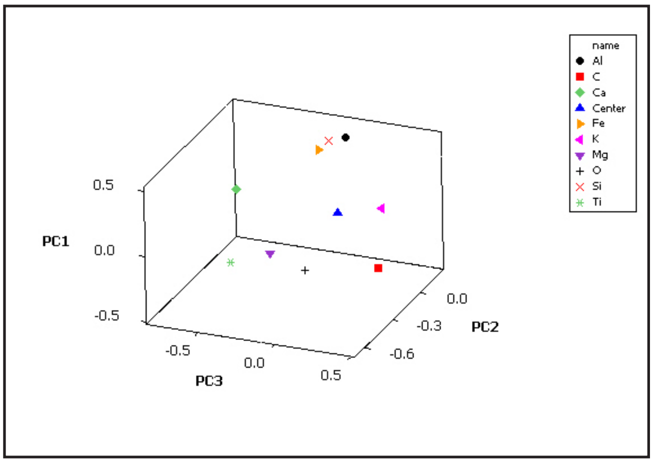
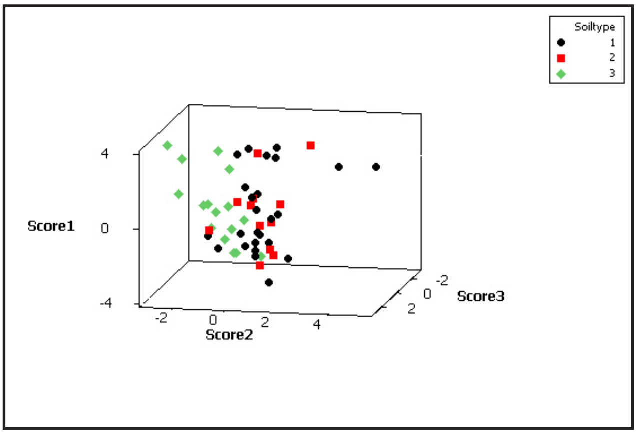
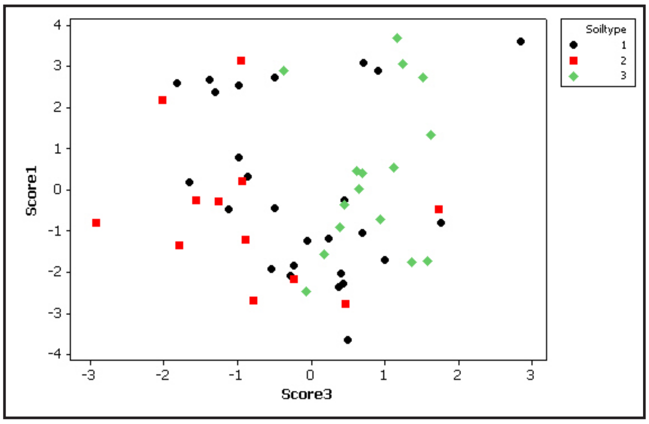


Figure 3. PCA loading 3-D plot of PC1 versus PC2 versus PC3



(a)



(b)

Figure 4. (a) Score plot of PC1 versus PC2 versus PC3 (b) Score plot of PC1 versus PC3 indicating the differentiation of soil samples according to their textures; 1= clay, 2= clay loam and 3= medium loam.

Acknowledgment

This work was supported by faculty of science, Silpakorn University, Thailand.

References

Black CA. Method of Soil Analysis, Part 2: Chemical and Microbiological Properties. Agronomy 9l. American Society of Agronomy. Wisconsin. 1965.

Boruvka L, Vacek O, Jehlicka J. Principal component analysis as a tool to indicative the origin of potentially toxic elements in soils. Geoderma 2005; 128: 289-300.

Dragovic S, Onjia A. Classification of soil samples according to their geographic origin using gamma-ray spectrometry and principle component analysis. Journal of Environmental Radioactivity 2006; 89: 150-58.

Gee GW, Bauder JW. Particle-size analysis. *In*: Methods of soil analysis part 1: physical and mineralogical methods (*Ed*: Klute A). Soil Science Society of America, Inc. and American Society of Agronomy, Inc. Madison, Wisconsin, USA. 1986; 383-410.

Jenkins R, Gould RW, Gedcke D. Quantitative x-ray spectrometry. Marcel Dekker Inc. 1995.

Jolliffe IT. Principal component analysis. Springer-verlag, Newyork. 1986.

Manzoor S, Shah MH, Shaheen N, Khalique A, Jaffar M. Multivariate analysis of trace metals in textile effluents in relation to soil and groundwater. Journal of Hazardous Material A 2006; 137: 31-37.

Mico C, Recatala L, Peris M, Sanchez J. Assessing heavy metal sources in agricultural soil of an European Mediteranean area by multivariate analysis. Chemoshere 2006; 65: 863-72.

McBride MB. Environmental chemistry of soils. Oxford University Press, Inc., New York, USA. 1994; 406 p.

Rencher AC. Methods of multivariate analysis. John wiley&son, Inc. 1995.

Sparks, DL, Page, AL, Helmke, PA, Loeppert, RH, Soluanpour, PN, Tabatabai, MA, Johnston CT, Sumner ME. Soil Science Society of America, Inc. and American Society of Agronomy, Inc. Madison, Wisconsin, USA. 1996.

Sposity G. 1989. The chemistry of soils. Oxford University Press, Inc. New York, USA. 277 p.

Stevenson FJ. Humus chemistry. John Wiley & Sons, Inc. New York, USA. 1994; 496 p.

Tariq SR, Shah MH, Shaheen N, Khalique A, Jaffar M. Multivariate analysis of selected metals in tannery effluents and related soil. Journal of Hazardous Material A 2005; 122: 17-22.

Received 4 May 2010
Accepted 12 June 2010

Correspondence to
Dr. Kamolchanok Panishkan
Faculty of Science,
Silpakorn University,
Nakorn Pathom, 73000
Thailand
Email: kamolcha@su.ac.th