Clustering Analysis of Soil Environmental Quality for Perennial Crop Recommendations in Vinh Long Province, Vietnam

Vo Quang Minh[1], Pham Thanh Vu[1*], Le Van Khoa[2], Tran Kim Tinh[2], Pham Cam Dang[1]

[1] Land Resources Department, Environment and Natural Resources College, Can Tho University, Can Tho, 90000, Vietnam

[2] Soil Science Faculty, College of Agriculture, Can Tho University, Can Tho, 90000, Vietnam

* Corresponding author's e-mail: ptvu@ctu.edu.vn

ABSTRACT

The study aimed to evaluate the soil environmental characteristics of Vinh Long Province's perennial crop-growing area using principal component analysis (PCA) and cluster analysis (CA). Soil environmental quality data were collected in eight districts of Vinh Long province for 27 physical and chemical parameters. CA and PCA analysis was used to group and identify critical parameters affecting perennial crops' soil environment. The findings demonstrate low to moderate soil compaction porosity, buffering capacity, and structure for perennial crops. In addition, the soil has a low pH, electrical conductivity, total soluble salts, aluminum, and cation exchange capacity. Although rich in nutrients, the content of organic matter, available phosphorus, cations, and trace elements is only low to moderate. CA results showed three districts suitable for strongly developing perennial crops: Tra On, Mang Thit, and Vung Liem. The PCA results showed that except for density, the buffer capacity of the soil, and dissolved Al^{3+}, the upcoming monitoring program must incorporate all remaining criteria. The study's findings offer crucial information to help the management organization devise strategies for enhancing and sustainably expanding perennial crops in the province. It is necessary to further evaluate the soil's environmental quality over time and soil depth and determine the frequency of monitoring in the study area.

Keywords: clustering, perennial crop, PCA, soil environmental quality.

INTRODUCTION

Due to its many applications, data mining and artificial intelligence communities are interested in data clustering. Without class information, it is challenging to identify the clustering scheme that best matches the raw partitions. When samples from the same group are comparable and grouped in clusters, clustering automatically separates the data points into the many groups the samples belong to. The data set has been divided into groups based on commonalities using a variety of clustering approaches that have been provided. These algorithms can generally be divided into viable clusters, soft (fuzzy), and complicated (sharp) methods [Höppner et al., 1999; Krishnapuram & Keller, 1993; Thompson et al., 1974].

Among various methods for choosing the essential variables (as components), principal component analysis (PCA) is one. It extracts low and high dimensions from the featured dataset to get as much data as possible. Also, visualization gains a lot more significance when variables are added. PCA is helpful when working with data with three dimensions or more. Asymmetric covariance or correlation analysis is performed. The underlying problem with multivariate statistics is that it is difficult to visualize data with several variables. Many variable groupings frequently move together in datasets with numerous variables. Multiple variables can measure the system's controlling principle. The abundance of usefulness makes measuring a wide range of characteristics possible. Information redundancy can be used to your advantage when it occurs. A new variable can be used instead of a group of existing variables to simplify the issue.

Agricultural land is one of the essential components of the ecosystem and a place to produce food for humans. According to Mulat et al. (2021), sustainable agriculture and production productivity are closely related to the soil environment quality. Therefore, soil environmental quality assessment has always been considered essential for managing and assessing soil property changes over space and time [Sulyman, 2018]. In addition, assessing soil environmental quality according to different land use purposes in areas with a tropical monsoon climate is very useful when making land use decisions and implementing sustainable land environmental management [Orobator, 2019]. Yet, because both natural factors (such as temperature, humidity, and rainfall) and human activities frequently change the qualities of the soil environment, it can be difficult to accurately determine the soil environmental quality (excessive fertilizers and pesticides) [Du et al., 2018; Saleh et al., 2021].

Vinh Long is a province located in the Mekong Delta, close to the lowest reaches of the Mekong River and positioned between the Tien and Hau Rivers. Compared to other regions of the Mekong Delta, this province had early and diverse agricultural growth. The province's total agricultural land area is 120,490.1 ha, accounting for 78.97% of the province's total natural land area. According to land use type, the province's agricultural land is divided into rice, perennial, annual, aquaculture, and other agricultural land. With nearly 50,000 hectares of land for perennial crops, Vinh Long has much potential to develop fruit trees such as oranges, pomelos, longans, and rambutans and bring high incomes to farmers in the province. However, in recent years, saline intrusion and excessive pesticide use have changed perennial crop soil properties [Department of Natural Resources and Environment of Vinh Long Province, 2020]. Therefore, to boost agricultural productivity in Vinh Long province in the Mekong Delta, Vietnam, and as a foundation for wise land use planning, it is required to constantly monitor and analyze the soil characteristics of the perennial fishing region. Principal component and cluster analysis are the most common multivariate statistical techniques for finding, categorizing, and minimizing the number of variables in soil environmental quality evaluation [Jiang et al., 2020; Tian et al., 2022]. Therefore, this study evaluated the soil environmental quality factor affecting perennial crops in Vinh Long province by clustering and the principle component analysis

method. The study's findings offer valuable scientific data that can help sustainably manage agriculture and increase agricultural output.

MATERIALS AND METHODS

Study area

The natural area of Vinh Long is 1,525.73 km², or 3.74% of the Mekong Delta's total area. The province of Vinh Long is located on a plain with a relatively flat landscape. Less than a 5° slope is present. The Tien, Co Chien and Hau Rivers run through the province. The subsurface, shaped like a basin, gradually rises along the Tien River, Hau River, Mang Thit River banks, and other significant rivers and canals from the province's center. Vinh Long is a tropical monsoon region with an annual average temperature of 27.4°C, an average humidity of 83%, and an average precipitation of 1,409 mm. It demonstrates the area's ideal climate for agricultural growth and development. The land in Vinh Long province has almost all been used up and has direct access to the sea. The province's four main types of soil resources are alluvial, sandy, and acid sulfate. Rice farming primarily uses 120,490.1 hectares of agricultural land (71,642.0 ha). Perennial crops, which comprise 46,676.5 hectares, or 38.74% of agricultural land, are planted primarily in fruit trees and are concentrated on islets between the Tien and Hau rivers and canals along the traffic axis in the shape of gardens and dwellings. Most are in the Long Ho, Tra On, and Vung Liem districts. The administrative map of the study area is shown in Figure 1.

Materials

Soil samples were collected in Vinh Long province's city and districts, including Vinh Long cities, Binh Minh town, Long Ho, Mang Thit, Vung Liem, Tam Binh, Binh Tan, and Tra On districts (Figure 1). The soil was analyzed with 27 parameters for agricultural production. The soil's physical properties include soil texture composition (sand, loamy sand, silt, clay), density, and buffering capacity. In addition, the chemical composition of the soil, including its acidity, organic matter (OM), total nitrogen (N), total phosphorus (P), available phosphorus (P_2O_5), soluble salts, pH_{H2O}, electrical conductivity (EC), and exchangeable cations (K^+, Na^+, Mg^{2+}, Ca^{2+}, Al^{3+}, Mn^{2+}, and Zn) are all essential factors.
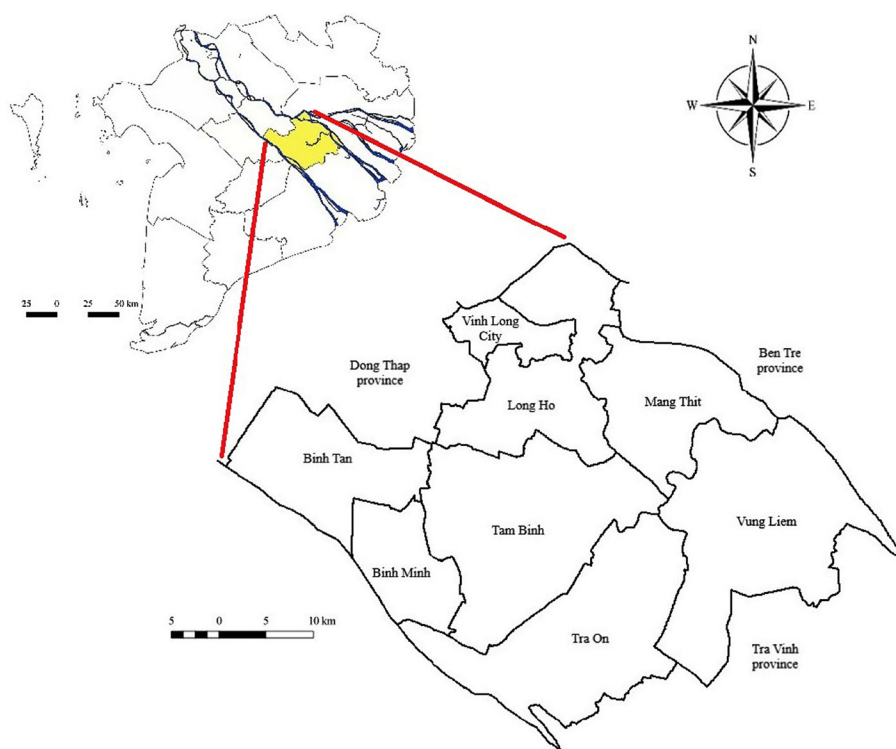
Figure 1. Map of Vinh Long and study area

Data analysis

Clustering analysis

Clustering is the process of grouping data objects into several groups with a high degree of similarity and a high degree of dissimilarity between them. Also, while things between groupings are rarely comparable, objects inside clusters have a high intraclass similarity. Thus, an effective clustering algorithm should yield results that have these characteristics (low interclass similarity).

This study applies clustering analysis (CA) to group soil sample survey locations according to physicochemical parameters. The CA results are displayed as a dendrogram, which groups the places with comparable soil environmental quality [Ibrahim, 2015].

The purpose of the clustering algorithm is to divide the given data set $X = \{x_1, x_2, …, x_n\}$ into groups $X_1, X_2, …, X_C$. The partition matrix U of size $c \times n$ may be represented as $U = [u_{ij}]$, $i = 1, 2, …, c$, and $j = 1, 2, …, n$, where u_{ij} is the membership of the sample x_j to cluster X_i. In the case of crisp partitioning, the following condition should be satisfied: $u_{ij} = 1$ if $x_j \in X_i$; otherwise, $u_{ij} = 0$ [Mulat et al., 2021].

Formally, the goal of clustering analysis can be represented as follows $\bigcup_{i=1}^{c} X_i = X,$

$$X_i \neq \emptyset, i = 1, 2, …, c,$$
$$X_i \cap X_j = 0, i = 1, 2, …, c, j = 1, 2, …, n. \quad (1)$$

In the case of fuzzy clustering, the membership u_{ij} is in interval [0,1]. A specific instance of fuzzy clustering may be seen in crisp clustering.

Principle component analysis

Conversely, principal component analysis (PCA) is used to identify the factors that significantly impact the research area's soil environment quality. Using PCA allows us to condense the representation space and retain only the parameters that offer discriminating information. The more significant the eigenvalue is in explaining the variability of the data, the larger it is, and it should ideally be greater than 1 [Mohab & Abdulaziz, 2022]. The correlation of load factors explains the data variables. With absolute values of load factor larger than 0.75, between 0.5 and 0.7, and less than 0.5, respectively, they are categorized into three levels: "strong," "moderate," and "weak" [Islam et al., 2017].

Even though various clustering analysis criteria lead to a range of clustering results, the following stages must typically always be followed by all clustering algorithms: a) *Data preparation*: data preprocessing; b) *Feature selection and extraction*: choose valuable attributes to preserve as

much information in the processed data as feasible; c) *Clustering*: based on the data structure and properties, choose the suitable clustering algorithm; d) *Clustering effectiveness evaluation*: to confirm the outcomes of the clustering, choose the suitable cluster effectiveness index., and e) *Results analysis*: The clustering results were examined with experimental data, and the proper conclusion was drawn.

The selection of the clustering technique and efficiency index will immediately impact the reliability and accuracy of the clustering results [Mahi et al., 2015]**.**

Primer V5.2 for Windows licensing software was used to analyze PCA and CA analyses (PRIMER-E Ltd., Plymouth, UK).

RESULTS AND DISCUSSION

Soil characteristics for perennial crops in the study area

Soil physical properties characterization

Table 1 shows the investigation findings for the soil's physical properties. The texture composition of soil samples in the study area is mainly loamy (0.002–0.02 mm), with an average value of 52.83%. Next is the clay particle size (<0.002 mm), which accounts for 41.72%, and the sand content (0.02–2 mm) in the soil is relatively low compared to the clay content of 5.46%. The soil sample's texture composition in the study area is mainly silty to clay soil, which is suitable for growing perennial crops. Soil with compacted surface bulk density at locations ranges from 1.1 to 1.3 g/cm³. The average value is 1.2 g/cm³. Otherwise, the average soil density in the study area reached 2.52 g/cm³. According to Nam et al. (2021), soils with a bulk density of less than 2.5 g/cm³ are assessed to have high humus content. However, if the soil bulk density is more significant than 1.2 g/cm³, it will lead to difficulties in

cultivation. Compacted soil limits water uptake and increases surface runoff and soil erosion, reducing soil nutrient content [Odey, 2018]. It shows that Long Ho, Mang Thit, Binh Tan, and Tra On districts may face more difficulties growing perennial crops than other districts.

The surface soil porosity in the study area had an average value of 48.74%, the highest in the Long Ho district (55.75%). Soil porosity should be above 50% because, according to Afriawan et al. (2021), it improves plant growth conditions and boosts soil respiration and microbial activity. Thus, only soil samples in Long Ho and Mang Thit districts have medium porosity, which is quite good for the growth of perennial plants. The remaining sites, on the other hand, all have soil porosities of less than 50%, indicating that the soil in these areas exhibits evidence of compaction and poor aeration and considerably impairs plant root absorption. Therefore, surface soil must often be tilled to promote soil porosity and aeration for plant root development in the research area and maximum nutrient absorption.

Soil chemical properties characterization

Table 2 shows the findings from examining chemical parameters in the soil. The soil pH of the study area ranges from 4.40 to 5.90 for pH_{H2O} and from 3.70 to 4.40 for pH_{KCl}. Low soil pH for perennial crops may be due to the depletion of cations during long cultivation, and the high use of inorganic fertilizers has increased H^+ ions in the soil [Mulat et al., 2021]. In addition, the growth of plant roots and the nutrient absorption mechanisms of plants also affect the release of H^+ ions into the soil [Dung et al., 2020]. According to Msimbira & Smith (2020), acidic soil would significantly impacts the growth and plant development and microorganisms in the soil. Otherwise, it will be difficult for plants to absorb the necessary minerals and nutrients, reducing crop yield. The EC value in the soil for perennial crops in the study area is relatively low, ranging from

Table 1. Soil physical properties

No.	Parameter		Mean	Standard deviation	Min	Max
1	Soil texture	Sand (%)	5.46	1.59	3.40	7.86
		Silt (%)	52.83	2.63	49.41	57.69
		Clay (%)	41.72	2.37	37.03	44.97
2	Bulk density	g/m³	1.20	0.07	1.10	1.30
3	Density	g/m³	2.52	0.08	2.40	2.60
4	Porosity	%	48.74	5.29	42.34	55.75

Table 2. Soil chemical properties

No.	Parameter		Unit	Mean	Standard deviation	Min	Max
1	pH	pH_{H2O}	-	5.19	0.60	4.40	5.90
		pH_{KCl}	-	4.01	0.21	3.70	4.40
2	EC	-	mS/cm	0.13	0.07	0.07	0.24
3	OM	-	%	3.08	0.42	2.50	3.60
4	Total soluble salts	-	%	0.05	0.03	0.01	0.12
5	CEC	pH7	meq/100 g	13.75	1.04	12.00	15.00
		pH8	meq/100 g	16.16	1.71	13.26	17.70
		Unbuffed	meq/100 g	12.95	3.83	7.33	18.00
6	Total N	-	%	0.18	0.03	0.13	0.22
7	Total P	-	%	0.17	0.04	0.11	0.22
8	Available phosphorus	-	mg/100 g	3.04	3.03	0.08	7.68
9	Total K	-	%	1.73	0.14	1.50	1.90
10	Total sulfur	-	%	0.23	0.08	0.11	0.37
11	Exchange Al^{3+}	-	meq/100 g	1.59	1.87	0.02	4.94
12	Dissolved Al^{3+}	-	meq/l	1.29	0.49	0.91	1.94
13	Cation exchange	K^+	meq/100 g	0.35	0.08	0.23	0.45
		Na^+	meq/100 g	0.52	0.17	0.27	0.79
		Ca^{2+}	meq/100 g	6.24	1.98	3.77	9.17
		Mg^{2+}	meq/100 g	3.27	1.10	1.20	4.53
14	Mn	-	mg/kg	46.73	28.45	21.88	93.70
15	Zn	-	mg/kg	28.28	35.45	6.12	90.93

0.07 to 0.24 mS/cm. Dung et al. (2020); and Nam et al. (2021) suggested that conductivity < 1 mS/cm will not affect growth and plant development.

The findings indicate that the soil samples had low to moderate organic matter content, with an average value of 3.08% and a range of 2.50% to 3.60%. However, the highest was in the Vung Liem district (3.60%), and the lowest was in the Binh Tan district (2.50%). Due to the high rate of soil organic matter oxidation and the comparatively long duration of perennial crops, the soil's relatively low organic matter content can be linked to both factors [Mulat et al., 2021]. Therefore, low soil organic matter content will reduce soil fertility and significantly affect crop yields [Gerke, 2022].

The concentration of soluble salt differs significantly between sampling locations in the study area, ranging from 0.01 to 0.12. The highest concentration of soluble salts was in Binh Minh Town (0.12%), while the lowest was in Vinh Long City (0.01%). According to Chau et al. (2021), the soil environment in the study area was assessed to be free of salinity (dissolved salt concentration < 0.128%) and had a negligible effect on the yield of the crops. In addition, according to the Department of Natural Resources and Environment of Vinh Long province (2020), it is said that people combine land work, shoveling, and pumping irrigation for crops in the dry season and during

salinization. Therefore, the accumulated salinity in the soil is not high, which improves crop yield.

According to Table 2, the pH 7 buffer CEC value, which ranged from 12 to 15 meq/100g, did not substantially differ between the sampling sites. Meanwhile, the buffered mean CEC value at pH 8.1 increased to an average of 16.16 meq/100g, and the unbuffered mean CEC value was 12.95 meq/100g. Soil development, organic matter concentration, and clay mineral composition significantly impact the soil CEC. [Dung et al., 2020]. The buffer capacity of the soil in the perennial crop area of Vinh Long province is low to moderate, with an average value of 229.04 meq OH. Better soil ecosystem resilience is associated with high soil buffer capacity [Dvoáková et al., 2022].

There were no significant differences in the study area's total N content between the sampling sites, ranging from 0.13% to 0.22%. According to research by Chau et al. (2021), soil samples in the perennial crop area of Vinh Long province had moderate to rich total N content. The soil's environmental quality in the study area ensures an adequate supply of nitrogen nutrients for plants. However, when the total N content is too high, it will lead to soil hardening and pollution, while too low total N will reduce soil fertility [Ma et al., 2022].

All soil samples were rich in total P content (0.11–0.22%), but available P content was low to

moderate and ranged from 0.08 mg/100g to 7.68 mg/100g. The comparatively good mineralization of organic materials may cause the soil's high total P content [Mulat et al., 2021]. The optimal P content for perennial plants, especially citrus, is 5–7 mg/100g [Loan et al., 2016]. Therefore, it can be seen that only soil samples in Tra On and Long Ho districts have available P content suitable for citrus needs. The results also showed that the available P content in soil samples with high total P content was also high. Therefore, some of the available P content in the soil was fixed. From there, providing available P for perennial crops to achieve high yields is necessary. Compared with the study of Dung et al. (2020), the amount of available P in the soil was lower (10–50 mg/100g) in the perennial crop region of Vinh Long province than it was in the pomelo growing region of Hau Giang province.

The research area's total K content ranged from 1.50% to 1.90% without significantly differing. However, according to Nam et al. (2021), K was found in moderate amounts overall in the study area. It is because perennial crops in the study area can meet their current crop needs and increase the overall K content of the soil. However, it is necessary to add potassium reserves to the soil for perennial crops, especially when the crop has a high demand for potassium. Therefore, the total K content was higher than that of the pomelo habitat (0.23%) in the study by Chau et al. (2021).

Sulfur is also necessary for growth and plant development. Sulfur is supplied to plants as an anion (SO_4^{2-}) and is readily leached out of the soil [Narayan et al., 2022]. Most of the sulfur in soil is present in organic matter and hence not accessible to the plants. Anionic form of sulfur shows that the total S content in soil samples is medium to high, ranging from 0.11 to 0.37% (Table 2). It shows that the total S content in the study area meets the demand for sulfur in crops. However, too high a soil sulfur concentration will also adversely affect plants through root damage, leaf burn, deformation, and reduced growth [Likus et al., 2018]. Therefore, each crop in the study region must maintain an appropriate total S concentration in the soil through proper management and fertilization regimes.

The exchange Al^{3+} content in the perennial crop area was recorded at a low level and had an average value of 1.59 meq/100g, while the low concentration of dissolved Al^{3+} had an average value of 1.29 meq/100g. Research results show that Al^{3+} content in soil is less likely to be toxic to perennial crops in the study area. Hung et al.

(2017) suggested that when soil pH is less than 4.5, Al^{3+} is more soluble. In addition, organic fertilizers will help reduce the amount of Al^{3+} exchanged and Al^{3+} dissolved in the soil through chelation [Dung et al., 2018].

The findings demonstrated that exchanged K^+ and Na^+ concentrations were low to moderate in soil samples from perennial crops, ranging from 0.23 to 0.45 meq/100g and 0.27 to 0.79 meq/100g, respectively. As can be observed, the amount of K^+ and Na^+ in the soil has little to no impact on how well plants grow and develop. Its outcome is comparable to Dung's study [Dung et al., 2020]. The amounts of K^+ and Na^+ in the soil used for cultivating pomelo in the Chau Thanh district of Hau Giang province were similarly found to be moderate (0.36–1.60 meq/100g and 0.2–0.7 meq/100g, respectively). Meanwhile, the study area's exchangeable Ca^{2+} and Mg^{2+} concentrations were moderately high, ranging from 3.77 to 9.17 meq/100g and 1.20 to 4.53 meq/100g, respectively.

Generally, soils in Mang Thit district have the lowest exchangeable cation content compared to other districts (0.23 meq/100g, 0.27 meq/100g, 3.77 meq/100g, and 1.20 meq/100g). Soils in Long Ho, Vung Liem, Binh Tan districts, and Vinh Long City had the highest K^+ (0.45 meq/100g), Na^+ (0.79 meq/100g), Ca^{2+} (9.17 meq/100g) and Mg^{2+} (4.53 meq/100g) content in the soil, respectively. Compared to the soil in the Dong Thap Muoi region in Hung's study [Hung et al., 2017], exchangeable Ca^{2+} and Mg^{2+} concentrations in the study area were much higher (1.59–1.69 meq/100g and 2.69–4.75 meq/100g).

On the other hand, the exchangeable K^+, Na^+, Ca^{2+}, and Mg^{2+} cations in the soil had an average value of 0.05 meq/100g, 0.51 meq/100g, 0.25 meq/100g and 0.16 meq/100g, respectively. Therefore, the content of exchangeable K^+, Ca^{2+}, and Mg^{2+} cations was low. At the same time, the Na^+ content in the soil was assessed as medium. This result is consistent with the study by Hieu et al. (2015). The exchangeable Na^+ content is always the highest and most highly variable among the four essential cations in the soil solution. Chaganti & Culman [2017] suggested that sodium is a substance that adversely affects the structure of the soil. On the contrary, calcium and magnesium facilitate soil cohesion and promote soil degradation. Therefore, reducing the exchangeable Na^+ content in the soil for perennial crops in the study area is necessary to ensure soil environmental quality and crop yield.

Finally, the analysis of some trace elements in the soil of perennial crops in Vinh Long province showed that Mn and Zn concentrations in soil samples ranged from low to moderate, with an average value of 46.73 mg/kg and 28.28 mg/kg, respectively. The highest concentrations of Mn and Zn were recorded in Tra On district (93.70 mg/kg) and the Mang Thit district (90.93 mg/kg). It may be because Zn is a potent antagonist of P. Therefore, when the soil is poor in P, the content of Zn in the soil increases, and vice versa. Besides, when the soil is too rich in P, it will reduce the ability to supply Zn to the plants and cause an imbalance of nutrients [Loan et al., 2016]. Therefore, adding enough Mn and Zn to the soil in the study area is necessary to avoid an imbalance of nutrients when the content is too low and toxicity to plants when the concentration is too high.

In Vinh Long province, the perennial crop soils often have silt to clay soil with relatively high density, porosity, buffering capacity, and low soil structure stability. It also typically has a relatively low soil pH. Low to moderate concentrations of OM and CEC, low Al^{3+} content, electrical conductivity, and total soluble salts indicate that the area has not been salinized and is not toxic to plants. The soil in the research area is also rich in N and P, although only low to moderate amounts of available P and moderate to high levels of total K and S are present. Base saturation, exchange cations, and trace elements are low to moderately prevalent. As a result, compared to other districts, the soil samples in the districts of Tra On, Mang Thit, and Vung Liem will be pretty favorable for vigorously developing perennial crops, notably citrus trees.

Clustering soil environmental quality in the study area

Cluster analysis using 27 physicochemical parameters of soil environmental quality in 2018 in 08 districts of Vinh Long province. At Euclidean distance 3, the soil sampling locations are divided into 6 clusters (Figure 2).

Clusters 1, 2, 3, and 5 have only one monitoring location: Binh Tan, Mang Thit, Long Ho, and Vinh Long City districts. Thus, the upcoming monitoring program must monitor these districts much more independently. Cluster 4 includes two districts, Vung Liem and Tra On. The main reason is that these two districts' physical and chemical characteristics are located in the eastern part of Vinh Long province. In addition, these two districts have the highest concentrations of OM, total P, available P, total K and S, and Al^{3+}, and the highest soil buffering capacity compared to the remaining districts.

Meanwhile, cluster 6 includes Binh Minh and Tam Binh districts due to similar soil environmental quality characteristics such as low total soluble salts, low nutrients, and trace elements, so they are grouped into one cluster. Therefore, Tra On and Tam Binh districts were selected based on soil characteristics and properties for future monitoring. Generally, the districts selected for long-term monitoring and representing soil environmental quality characteristics in Vinh Long province are Tra On, Tam Binh, Binh Tan, Mang Thit, Long Ho, and Vinh Long city districts.

Principal components are a new collection of variables the PCA approach can provide [Shihab et al., 2020]. The original variables are combined linearly to form each primary component. As all
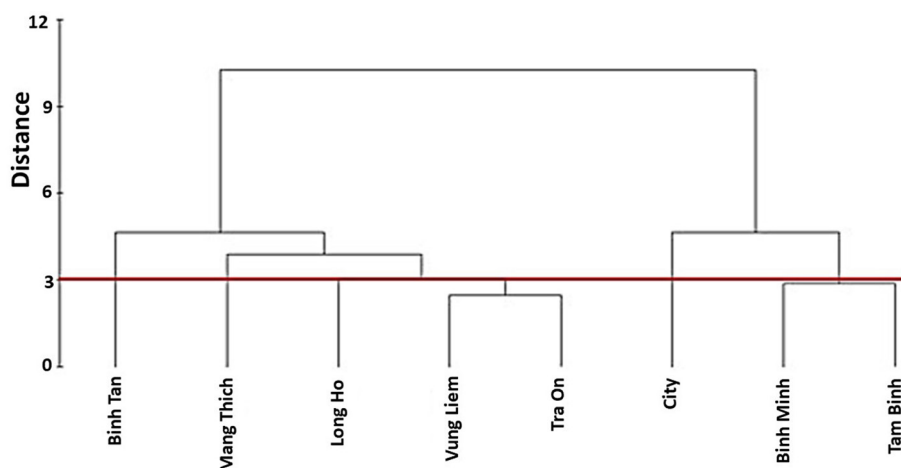

Figure 2. Results of soil environmental quality classification

significant components are orthogonal, there is no redundant data, laying the foundation for the orthogonal data space. A single axis in a matrix makes up the first principal component. The most viable option available for the first axis is the variable variance. Another axis in a matrix perpendicular to the first is related to the second primary component. The observation on this axis creates an additional new variable. Once more, this variable's variance is the highest of all the options available for the second axis.

Table 3 presents the principal component analysis (PCA) results and identifies 7 factors that explain 100% of the variation in soil environmental quality. PC1 to PC6 are the main factors, explaining 97.3% of the variation with rates of 36.4%, 19.6%, 14.8%, 11.9%, 10.2%, and 4.4%, respectively. These PCs have eigenvalue values greater than 1, contributing significantly to the variation in the data. The PC7 is only regarded as a secondary source and does not substantially impact the variation in soil environmental quality in the research area.

The contribution of physical and chemically connected soil environmental quality factors, including density, porosity, and cation exchange capacity, explains PC1 (CEC buffered at pH 8.1). Like PC1, PC2 through PC5 are influenced by physical traits and mainly contribute via the cations exchange and Al^{3+}, total soluble salts, Zn, and other elements. Soil nutrients (N, P, and K total, total S, easily digestible phosphorus) significantly affect soil environmental quality. In the meantime, Mn, unbuffered CEC, and saturated cations in the soil are the main contributors to PC6. The results of PCA analysis show that the sources affecting soil environmental quality in the perennial crop area of Vinh Long province are related to natural processes (pH_{KCl}, pH_{H2O}, EC, density, porosity, organic composition world, CEC, exchangeable Al^{3+}, total soluble salts) and related sources of nutrients (OM, total N, P, K, total S, readily digestible phosphorus, exchangeable and saturated cations, Mn, Zn). This outcome aligns with Abdel-Fattah et al. (2021), suggesting that EC, organic matter, CEC, NPK, and soil mechanical composition significantly influence soil environmental quality.

Table 3. Key parameters influencing soil environmental quality

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Soil texture	0.142	0.095	0.137	-0.210	-0.490	0.022	0.026
Buld density	-0.242	-0.119	0.292	-0.088	-0.087	0.276	0.039
Density	-0.328	-0.091	0.058	0.044	0.019	-0.043	-0.001
Porosity	-0.329	-0.085	0.041	-0.032	-0.021	-0.076	-0.045
pH_{KCl}	-0.028	-0.362	0.070	0.339	-0.048	-0.053	-0.174
OM	-0.092	0.418	0.078	0.036	0.143	-0.062	0.192
CEC at pH 7	0.031	-0.044	0.432	0.298	-0.126	-0.132	0.105
Total N	-0.007	0.396	0.173	0.176	0.093	0.009	0.27
Total P	-0.063	0.329	-0.008	0.315	-0.213	0.126	-0.259
Total K	-0.236	0.074	-0.325	0.010	0.106	0.276	-0.086
Total S	-0.015	0.365	-0.044	0.234	0.207	-0.289	0.156
Total dissolved salts	0.083	0.227	0.319	0.109	-0.056	-0.107	-0.662
pH_{H2O}	0.183	-0.215	-0.043	0.353	0.176	0.211	0.047
EC	-0.146	0.008	-0.376	0.067	-0.304	-0.247	-0.11
Available phosphorus	-0.187	0.096	-0.115	0.092	-0.431	0.274	0.266
Cation exchange	0.196	-0.211	-0.037	0.378	-0.112	-0.077	-0.072
Al^{3+} exchangeable	0.043	0.015	-0.478	0.154	0.188	-0.081	-0.130
Mn	-0.239	-0.076	0.087	0.277	0.098	0.436	0.072
Zn	-0.162	0.015	0.161	-0.283	0.396	0.082	-0.296
CEC at pH 8.1	-0.318	-0.128	0.067	-0.033	0.007	-0.140	0.008
CEC unbuffered	-0.271	-0.198	0.114	-0.012	0.073	-0.312	0.128
Al^{3+} Soluble	-0.279	0.079	-0.122	-0.104	-0.262	-0.123	-0.197
Eigenvalues	8.73	4.70	3.55	2.87	2.45	1.05	0.65
%Variation	36.40	19.60	14.80	11.90	10.20	4.40	2.70
Cum. %Variation	36.40	56.00	70.80	82.70	92.90	97.30	100.00

CONCLUSIONS

The soil pH is low, the soil texture is primarily silty-clay, and the soil tends to be compacted by low porosity, according to the soil environmental quality characterization results of the perennial crop region of Vinh Long Province. The EC content, total soluble salts, Al^{3+}, and cation exchange capacity (CEC) were low. The buffer capacity is modest to moderate. While the soil's NPK concentration ranged from moderate to rich, its organic matter content was low to moderate. Low to moderate concentrations of cations, trace elements, and available phosphorus can impact plant growth. Soil samples in Tra On, Mang Thit, and Vung Liem districts are suitable for properly growing perennial crops.

CA analysis results suggest the Tra On, Tam Binh, Binh Tan, Mang Thit, Long Ho, and Vinh Long districts in the soil environmental quality monitoring program still represent the soil environmental characteristics of Vinh Long province. The results of the PCA analysis showed that there were 06 main PCs that explained 97.3% of the change in soil environmental quality in the study area. Key parameters that need to be continuously monitored include pH_{KCl}, pH_{H2O}, EC, density, porosity, texture composition, CEC, Al^{3+} exchange, total dissolved salts, organic matter, a total of N, P, K, and S, available phosphorus, exchange and saturation cations, Mn, and Zn. Further studies must focus on assessing soil environmental characteristics over time and soil depth while determining the frequency of environmental monitoring of perennial crops in the study area.

Acknowledgments

REFERENCES

1. Abdel-Fattah, M.K., Mohamed, E.S., Wagdi, E.M., Shahin, S.A., Aldosari, A.A., Lasaponara, R., & Alnaimy, M.A. 2021. Quantitative evaluation of soil quality using principal component analysis: The case study of El-Fayoum Depression Egypt. Sustainability, 13(4), Article 4. https://doi.org/10.3390/su13041824

2. Afriawan A., Ahmad A., and Gusli S. 2021. Analysis of plant root properties, texture, and porosity that affects landslides in Tangka Sub-Watershed. IOP Conf. Ser. Earth Environ. Sci., 886(1), 012093. doi: 10.1088/1755-1315/886/1/012093.

3. Chaganti V.N. and Culman S.W. 2017. Historical perspective of soil balancing theory and identifying knowledge gaps: A review. Crop Forage Amp Turfgrass Manag., 3(1) (Accessed: 24.3.2023). https://www.cabdirect.org/cabdirect/abstract/20193283156

4. Chau N.N.B., Truyen D.M., Dan T.H., and Loi L.V. 2021. Assessment of difference on land environment indicator in different landscapes in Cu Lao Dung area. Dong Thap Univ. J. Sci., 10(3), 56–63.

5. Department of Natural Resources and Environment of Vinh Long province. 2020. Announcement of results of investigation and assessment of soil quality and land potential for the first time in Vinh Long province. https://vinhlong.gov.vn/stnmt/xem-chi-tiet-tin-tuc/id/167393

6. Du T.T., Mi N. T.H., Minh V.Q., and Khoa L.V., 2018. Application of GIS in the mapping of soil fertility distribution in the Mekong Delta. Int. Symp. Geoinformatics Spat. Infrastruct. Dev. Earth Allied Sci. 60–65.

7. Dung L.V., Guong V.T., Linh N.D., and Thu T.A. 2018. Improving saline acid sulfate soils and rice yield by using organic amendment and lime in green house condition. Can Tho Univ. J. Sci., vol. 54, Can Tho University. doi: 10.22144/ctu.jsi.2018.067.

8. Dung T.V., Dang L., Nguyen Van Q., Toan L., and Hung N.N. 2020. Morphological and physico-chemical properties of the raised-bed soils cultivated with 5 Roi pomelo in Chau Thanh district - Hau Giang province. Can Tho Univ. J. Sci., vol. 56, 130–137.

9. Dvořáčková H., Dvořáček J., Hueso González P., and Vlček V. 2022. Effect of different soil amendments on soil buffering capacity. PLoS ONE, 17(2), e0263456. doi: 10.1371/journal.pone.0263456.

10. Gerke J. 2022. The Central Role of Soil Organic Matter in Soil Fertility and Carbon Storage. Soil Syst., 6(2). doi: 10.3390/soilsystems6020033.

11. Hieu N.T., Tinh T.K., and Thành V.C. 2015. Chemical characteristics of saline soil where "Ski" variety of rice was grown in Hong Dan-Bac Lieu. Can Tho Univ. J. Sci., 2(40), Art. no. 40.

12. Höppner F., Klawonn F., Kruse R., and Thomas R. 1999. Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition. Wiley. 1999. Accessed: Mar. 24, 2023. https://www.wiley.com/en-us/fuzzy+cluster+analysis%3a+methods+for+classification%2c+data+analysis+and+image+recognition-p-9780471988649

13. Hung T.V., Toan L.P., Dung T.V., and Hung N.N. 2017. Morphological and physio-chemical properties of acid sulfate soils in Dong Thap Muoi. Can Tho Univ. J. Sci., 2. 1–10. doi: 10.22144/ctu.jsi.2017.047.

14. Ibrahim A. 2015. Evaluation of surface water quality using multivariate techniques in terengganu river basin. International Journal of Development Research (IJDR). Int. J. Dev. Res., 5(2), 3421–3427.

15. Islam M., Lenz O., Azad A., Ara H., Rahman M., and Hassan N. 2017. Assessment of Spatio-Temporal Variations in Water Quality of Shailmari River, Khulna (Bangladesh) Using Multivariate Statistical Techniques. J. Geosci. Environ. Prot., 5, 1–26. doi: 10.4236/gep.2017.51001.

16. Jiang M., Xu L., Chen X., Zhu H., and Fan H. 2020. Soil Quality Assessment Based on a Minimum Data Set: A Case Study of a County in the Typical River Delta Wetlands. Sustainability, 12(21), Art. no. 21. doi: 10.3390/su12219033.

17. Khoa L. V. and Ti N. 2013. Soil stability classification and factors influencing to the soil structural stability of alluviral soils in the Mekong Delta, Vietnam. Can Tho Univ. J. Sci., 26, no. B,. 219–226.

18. Krishnapuram R. and Keller J.M. 1993. A possibilistic approach to clustering. IEEE Trans. Fuzzy Syst., 1(2), 98–110, doi: 10.1109/91.227387.

19. Likus J., Pietrzykowski M., and Chodak M. 2018. Chemistry of Sulfur-Contaminated Soil Substrate from a Former Frasch Extraction Method Sulfur Mine Leachate with Various Forms of Litter in a Controlled Experiment. Water. Air. Soil Pollut., 229. doi: 10.1007/s11270-018-3716-2.

20. Loan N.T.P., Thao L.M., Minh L.C.T., and Tuan N.T. 2016. Investigation on soil properties in orange growing gardens in Cao Phong town, Hoa Binh province, Vietnam Soil Sci., 47, 16–21.

21. Ma. J., Cheng, J., Wang, J., Pan, R., He, F., Yan, L., & Xiao, J. 2022. Rapid detection of total nitrogen content in soil based on hyperspectral technology, Inf. Process. Agric., 9(4), 566–574. doi: 10.1016/j.inpa.2021.06.005.

22. Mahi H., Farhi N., and Labed K. 2015. Remotely Sensed Data Clustering Using K-Harmonic Means Algorithm and Cluster Validity Index, presented at the IFIP Advances in Information and Communication Technology. 105–116. doi: 10.1007/978-3-319-19578-0_9.

23. Mohab A.K. and Abdulaziz I.A. 2022. Assessment of Physicochemical Water Quality using Principal Component Analysis: A Case Study Wadi Hanifa, Riyadh. RiyadhCivil Eng Res J, 12(5). doi: DOI 10.19080/CERJ.2022.12.555850.

24. Msimbira, L., & Smith, D. 2020. The Roles of Plant Growth Promoting Microbes in Enhancing Plant Tolerance to Acidity and Alkalinity Stresses. Frontiers in Sustainable Food Systems, 4. https://doi.org/10.3389/fsufs.2020.00106

25. Mulat Y., Tsehai K., Bedadi B., and Yusuf M., 2021. Soil quality evaluation under different land use types in Kersa sub-watershed, eastern Ethiopia. Environ. Syst. Res., 10. doi: 10.1186/s40068-021-00224-6.

26. Nam T.S., Khanh H.C., Thao H.V., & Thuan N.C. 2021. Physical and chemical characteristics of soil inside and outside the full-dyke systems in Phu Tan district, An Giang province. Can Tho University Journal of Science, 57, Article Special issue on Environment and climate change. https://doi.org/10.22144/ctu.jsi.2021.033

27. Narayan O., Kumar P., Yadav B., Dua M., and Johri A., 2022. Sulfur nutrition and its role in plant growth and development. Plant Signal. Behav. doi: 10.1080/15592324.2022.2030082.

28. Odey S., 2018. Overview of Engineering Problems of Soil Compaction and Their Effects on Growth and Yields of Crops. doi: 10.13140/RG.2.2.11522.22722.

29. Orobator P.O., 2019. Investigating soil quality indicators response to perennial cropping systems in lowland areas of Northern Edo State, Nigeria. Int. of Social Sci. (Accessed: 24.03.2023) https://www.academia.edu/43642719/Investigating_Soil_Quality_Indicators_Response_to_Perennial_Cropping_Systems_in_Lowland_areas_of_Northern_Edo_State_Nigeria

30. Saleh A.M., Elsharkawy M.M., AbdelRahman M.A.E., and Arafat S.M., 2021. Evaluation of Soil Quality in Arid Western Fringes of the Nile Delta for Sustainable Agriculture. Appl. Environ. Soil Sci., vol. 2021. 1–17. doi: 10.1155/2021/1434692.

31. Shihab A., Dawood F., and Kashmar D., 2020. Data Analysis and Classification of Autism Spectrum Disorder Using Principal Component Analysis. Adv. Bioinforma., vol. 2020, 8 pages. doi: 10.1155/2020/3407907.

32. Sulyman M., 2018. Response of soil properties to land use change from wheat to pig production, J. Pollut. Eff. Control, 6. doi: 10.4172/2375-4397.1000211.

33. Thompson M., Duda R.O., and Hart P.E., 1974. Pattern classification and scene analysis in Leonardo. p. 370. doi: 10.2307/1573081.

34. Tian Y., Xu Z.,Wang J., and Wang Z., 2022. Evaluation of soil quality for different types of land use based on minimum dataset in the typical desert Steppe in Ningxia, China. J. Adv. Transp., vol. 2022. 1–14. doi: 10.1155/2022/7506189.