Prepublished copy

Nature Environment & Pollution Technology

This is a peer-reviewed prepublished version of the paper

Type of the Paper (Original Research)

Assessing Soil Health Through Multivariate Analysis: A Focus on Durian Cultivation in Cho Lach, Ben Tre Province, Vietnam

Nguyen Van Phuong†

Institute of Environmental Science, Engineering & Management, Industrial University of Ho Chi Minh City (70000), Vietnam

†Corresponding author: Nguyen Van Phuong; nguyenvanphuong@iuh.edu.vn

ORCID IDs of Authors: 0000-0001-7709-3880

Key Words	Soil quality, Soil quality index (SQI), PCA, Minimum data set (MDS), Durian growing area
DOI	https://doi.org/10.46488/NEPT.2025.v24i04.D1767 (DOI will be active only after the final publication of the paper)
Citation for the Paper	Phuong, N. V., 2025. Assessing soil health through multivariate analysis: A focus on durian cultivation in Cho Lach, Ben Tre Province, Vietnam. *Nature Environment and Pollution Technology*, 24(4), p. D1767. https://doi.org/10.46488/NEPT.2025.v24i04.D1767

ABSTRACT

Monitoring and evaluating soil quality is a trend in precision farming and sustainable agricultural management. This study used a multivariate analysis to evaluate soil quality in durian-growing areas in Ben Tre, Vietnam. Twelve representative composite soil samples were collected, and nine selected soil indexes were determined, including pH, EC, TOC, Bulk density, available phosphorus, NH4+, CEC, clay content, humus content, and water-holding capacity. The dataset was transformed into new variables using principal component analysis (PCA), deriving relative weights (Wi) and soil normalization scores (Si), which were subsequently utilized to determine the soil quality index (SQI). The results of the study were to identify the MDS set consisting of 3 principal components that completely explained 84.33% of the variance of the dataset. The 3 indicators (including % clay, EC, and available phosphorus) represented the principal components. The current SQI of the study area was mostly at the average level (accounting for 83.3%). The results of SQI calculation based on PCA can help save time, reduce laboratory work costs, and support precise and efficient agricultural management.

1. INTRODUCTION

Agricultural land is degraded by a combination of internal processes (climate change/weathering, erosion, sedimentation, and geology) and external processes due to human activities (land management strategies, waste management, erosion, deforestation in agriculture, urban planning, and use of agrochemicals) (Damiba et al. 2024).

This land degradation poses major threats to the environment, productivity, and sustainability of agricultural activities (Damiba et al. 2024),

Several methods have been developed to quantify soil quality (SQ) that have been used as decision support tools. Despite various methods, none have gained widespread adoption or recognition because of the intricate and heterogeneous nature of soil systems. This variability often results in inconsistent outcomes when evaluating the same geographical area. (Damiba et al. 2024)

Most methods use physical and chemical indicators and separate or ignore biological indicators. Using this method requires knowledge, and the selection of representative indicators is mostly based on the judgment of "experts". Finally, due to the intensive field and laboratory work, cost, and time constraints of being able to process a large set of indicators, it is necessary to reduce a large data set to a smaller set. (Damiba et al. 2024). Nowadays, the use of multivariate techniques of principal component analysis (PCA) for statistical data reduction has become more popular. (Abdu et al. 2023). Studies have confirmed that using PCA to reduce the data set is more sensitive and better in soil quality assessment (Damiba et al. 2024; Rajput et al. 2023).

In the Mekong Delta of Vietnam, agricultural activities such as intensive farming, perennial crop cultivation, and excessive use of fertilizers have led to soil degradation. This has resulted in declining soil quality and crop productivity. Therefore, urgent management solutions are needed to mitigate risks to agriculture, one of which is soil quality control based on soil quality index assessments. Durian, currently an emerging agricultural product in Vietnam, holds a very high export value, reaching 3.3 billion USD, accounting for 50% of the total agricultural export value in Vietnam in 2024, according to the Vietnam Customs General Department.

However, studies using PCA to assess soil quality in Vietnam, especially in durian-growing areas in the Mekong Delta, are still lacking (Table 1). The main parameters affecting soil quality, derived from an extensive dataset (LDS), include nine factors: bulk density, clay content, silt content, water holding capacity, pH, EC, TOC, available phosphorus, and NH_4^+, assessed using PCA for durian cultivation in Ben Tre, Vietnam. The findings of the study provide farmers, managers, and policymakers with a straightforward, efficient, and cost-effective method to enhance agricultural practices.

Table 1 Some studies assessing soil quality in Vietnam

No	Location	Evaluation parameters and methods	References
1	Mangrove forests of Thai Binh province	pH, heavy metals and nutrients, and PCA method	(Nguyen et al. 2024)
2	Tram Chim National Park, Dong Thap Province, Vietnam	pH, total nitrogen, total phosphorus, total acidity, organic matter, and total exchangeable iron and aluminum. Cluster analysis and PCA analysis	(Giao et al. 2023)
3	Agricultural region of A Luoi district, Central Vietnam	Differences in soil organic carbon, soil total nitrogen, and soil pH. Using mean and difference comparison method by ANOVA	(Pham et al. 2018)
4	Rice and corn farming systems in the mountainous regions of Central Vietnam	Particle size distribution, pH, organic carbon, total nitrogen, total phosphorus, available phosphorus, CEC (phương pháp so sánh)	(Van Binh et al. 2013)

2 Materials and research methods

2.1 Field sampling method

60 soil samples were collected from 12 locations in durian-growing areas of Cho Lach district, Ben Tre province, Vietnam, in January 2024 (Figure 1). Two major durian-growing areas were selected, with 5 samples from area 1 and 7 from area 2, representing homogeneous fields based on tree age and growth. Samples were taken at 0-30 cm depth as this surface layer is nutrient-rich, biologically active, and highly sensitive to environmental changes. Each composite sample was collected from a 10 m diameter area, mixing soil from 4 corners and the center. Samples were air-dried, crushed, and sieved (2 mm) for analysis. Indicators used in assessing soil quality included physical properties

(density, clay, silt, sand) and chemical/nutritional properties (pH, EC, TOC, NH4+, available P). Determination analysis: Bulk density (TCVN 8305:2009), TOC (by Walkley Black method), pH, and EC (ISO 10390:1993), available P (ISO 11464, and TCVN 5256:2009), NH4+ (Baethgen & Alley. 1989), and soil texture (Bouyoucos. 1962).

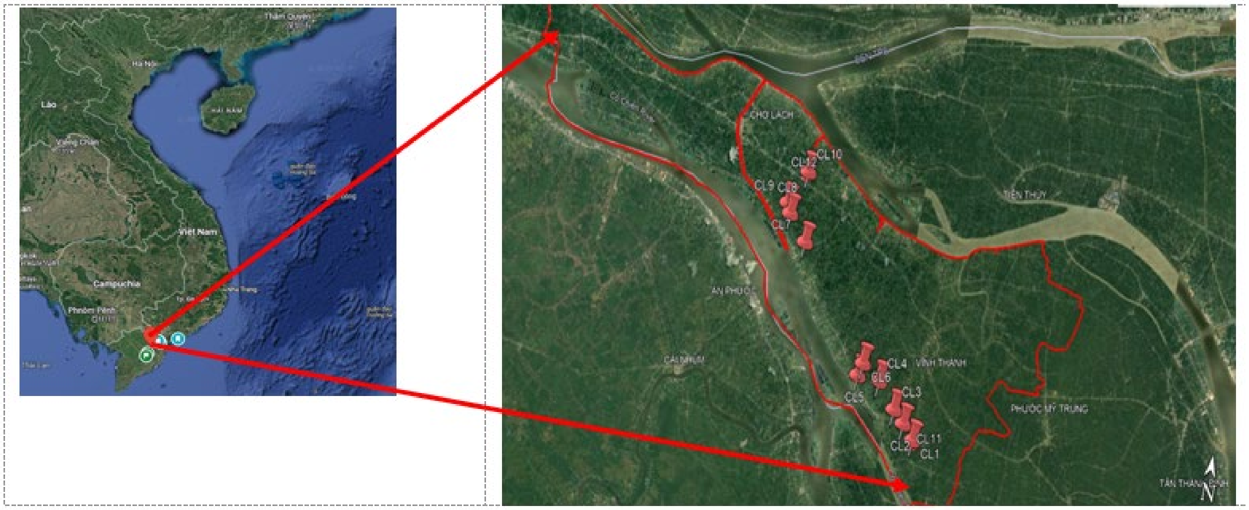

Figure 1 Location map of Cho Lach district, Ben Tre province, Vietnam (Source: Google Maps)

2.2 *Soil Quality Index (SQI) Assessment*

To determine the influential soil quality indicators in the study area, a statistical analysis on Excel and SPSS 23 was conducted. A Pearson correlation matrix was constructed to find out the degree of correlation between the study variables. After finding the correlated variables, principal component analysis (PCA) was developed. (Rangel-Peraza et al. 2017)

By employing PCA, the dataset's dimensionality is minimized through the extraction of principal components (PCs) and the analysis of orthogonal variable correlations, streamlining the data structure. Furthermore, PCA converts a large data set with correlated variables into a set of uncorrelated indices. PCA involves the following steps: (i) normalizing the variables, (ii) establishing a correlation matrix, (iii) determining PCs with eigenvalues, and percentage of variance, (iv) removing PCs with smaller eigenvalues (eigenvalues <1), and (v) establishing a PC matrix with influencing factor loadings. The communalities are calculated as the percentage of variance accounted for by each variable in the PC. Principal components (PCs) with eigenvalues exceeding 1 and explaining a minimum of 30% of the data variability were chosen for further analysis. Principal components should contribute >70% of the data variation (Salem & Hussein. 2019). Removing indicators with loadings below 0.3 in PCA ensures that the model focuses on variables with strong influence, reduces noise, simplifies results, and increases the explanatory power of principal components. From these PCs, only variables with significant loadings were included in the minimum dataset (MDS). These "highly loaded" variables were identified as those with the greatest weight on a specific PC, along with others whose absolute loadings fell within 30% of the highest recorded values. (Abdu et al. 2023). When several indicators are identified within the same principal component (PC), the Pearson correlation matrix is employed to assess their significant relationships ($p < 0.05$) and eliminate redundancy. Indicators displaying the highest factor loadings are prioritized if they exhibit strong correlations ($r > 0.5$). In cases where correlations are weak ($r < 0.5$), all indicators are retained for further consideration (Damiba et al. 2024)

2.2.1 *Normalization of Indicators*

The indicators measured vary in scale and units, necessitating their transformation into standardized scores ranging from 0 to 1. This normalization process enables the integration and averaging of diverse indicators into a unified value, facilitating the assessment of soil functions and processes. Additionally, it ensures that critical data points are not overlooked during evaluation. Both linear and non-linear scoring methods are employed to convert these indicators into dimensionless units within the 0–1 range (Damiba et al. 2024)

Once the soil quality indices have been analyzed, the interpretation of the values of the selected influencing parameters is clearly defined. Without an interpretation system, the indices cannot be used in practice. An advanced approach to standardizing soil quality indices is to establish standard non-linear scoring functions, usually of the form i) more is better (Formula 1), ii) optimal range (Formula 3), or iii) less is better (Formula 2), which are the most common in soil science. The shape of such curves is established based on a combination of reference values and expert judgment. (Bünemann et al. 2018).

In the "more is better" approach, the value of each observation is normalized by dividing it by the maximum observed value, ensuring the highest value scores 1 and all others are scaled proportionally below 1, as outlined in formula (1) (Bandyopadhyay & Maiti. 2021).

$$S_i = LSF = \frac{X}{X_{max}} \tag{1}$$

In scenarios where "less is better," the minimum observed value is divided by each observation, assigning a score of 1 to the lowest value and scores less than 1 to all other values, as described in formula (2):

$$S_i = LSF = \frac{X_{min}}{X} \tag{2}$$

Here, Si represents the linear scoring function (LSF), which ranges between 0 and 1, where X is the measured value of a specific soil parameter, and Xmax and Xmin denote the maximum and minimum observed values for that parameter, respectively. More is better, including moisture, clay, humus, N, and P; Less is better, including bulk density (Damiba et al. 2024)

For the non-linear scoring function, the soil indices are transformed according to the sigmoidal curve equation. formula (3) is as follows:

$$S_i = NLSF = \frac{a}{\left[1 + \left(\frac{X_i}{X_{imean}}\right)^b\right]} \tag{3}$$

The non-linear scoring function (NLSF) operates on a scale from 0 to 1, with its peak value fixed at 1. In this context, Xi denotes the measured value of soil index i, and Ximean represents the average value of the same index. The slope parameter, b, is set to -2.5 for scenarios where higher values are preferable ("more is better") and +2.5 for cases where lower values are desired ("less is better"). pH and EC have optimum thresholds (Damiba et al. 2024)

2.2.2 Calculation of soil quality index (SQI)

SQI is calculated from formula (4):

$$W - SQI = \sum_{i=1}^{n} W_i . S_i \tag{4}$$

Where: Si is the score of index i, n is the total number of relevant indices (Damiba et al. 2024)

Wi is the weight coefficient of index i calculated according to Formula (5)

$$W_i = \frac{PC_i}{\sum PC} \tag{5}$$

PCi is the loading of principal component i and ΣPC is the sum of the loadings of principal components with eigenvalues >1. Si is calculated based on the formulas in normalizing soil quality indices. Finally, SQI is calculated. The classification criteria for the main SQI indices include: very low (0-0.19); low (0.20-0.39); medium (0.4-0.59); good (0.6-0.79) and very good (0.8-0.99). (Damiba et al. 2024)

2.3 Data processing

To identify the soil quality indices that have a greater influence in the study area, a statistical analysis was conducted on Excel and SPSS 23. Initially, a descriptive statistical analysis was conducted to examine variability of the indices, identifying any outliers or unusual data points. Following this, a Pearson correlation matrix was developed to assess the strength and direction of relationships among the variables under investigation. PCA reduced the data set and constructed linear combinations (principal components) of the original variables that explained most of the total original variation. Pearson correlation coefficient determined the correlation between the investigated soil properties, while PCA was used to select MDS and to determine SQI.

3 RESULTS AND DISCUSSION

3.1 Determination of soil quality physicochemical parameters

The results of 9 selected physical and chemical soil indicators of 12 combined soil samples of the orange growing area of Ben Tre, Vietnam, are presented in detail in Table 2. The results showed that:

The bulk density of 12 soil samples in the growing area has an average value of 0.98 g cm^{-3}, ranging from (0.90-1.06), in which CL6 is the highest at 1.06 g cm^{-3} and the lowest at CL 1 is 0.90 g cm^{-3}, Table 2. According to field observations, sample CL1 is the soil for growing trees in the commercial stage, using more foliar fertilizers, so the air permeability is still maintained high and more porous. The overall bulk density of the study area is the only one with sample CL 1 (lowest) and CL 6 (highest) that is different from the other samples; the remaining samples have no significant difference (according to One-way ANOVA). The results show that the partition is not large. This can be explained by the fact that currently the value of durian is very high, so gardeners always try to use a lot of organic amendments to maintain the physical properties of the soil. The results are similar to previous studies when determining the bulk density of the Mekong Delta orange growing area fluctuates from 0.71 = 1.09 g cm^{-3} (Phuong. 2024).

The average pH was 4.67, the lowest was 3.99, and the highest was 5.42 for samples CL4 and CL1, respectively. The pH values of these two samples were significantly different from the remaining 10 samples (according to One-way ANOVA analysis), with pH ranging from 4.2 to 5.0. The overall pH of the study area formed 6 statistically significantly different areas in increasing order, including CL 4; (CL7, CL2, CL5); (CL9, CL10); (CL8, CL11, CL6); (CL12, CL3) and CL1. This shows that the pH fluctuation is very large, which may be due to the nature of the soil properties of the growing area (the location of the soil samples was collected with its characteristics according to field observations) and may also be due to the cultivation process such as tillage, fertilization, and use of pH improving substances. According to the assessment in the study of Amacher et al.., the soil pH in the study area is moderately acidic. Acid-intolerant crops will be affected depending on the level of Al and Fe leaching (Amacher et al. 2007). Because durian trees only grow well in soil with a pH of 5.5 to 6.5 (Amran et al. 2023), durian gardeners in the Mekong Delta always monitor the pH throughout the cultivation process. The results of the pH analysis of the soil samples were similar to previous studies in the orange-growing area of the Mekong Delta, with a pH ranging from 3.65 to 6.8. (Phuong. 2024). This may be because the durian growing area in the study also originated from acid soil, a typical soil type in the Mekong Delta (Husson et al. 2000).

The EC of the soil samples had an average value of 0.08, the lowest was 0.05, and the highest was 0.11 mS $cm^{-1,}$ corresponding to samples (CL1, CL11) and (CL4, CL7) (Table 2). The overall EC of the study area formed 5 statistically significant areas from low to high, including (CL1, CL11); (CL3, CL12); (CL6, CL2, CL5); (CL8, CL9, CL10) and (CL4, CL7). This shows that, similar to pH, EC values fluctuate greatly. Excluding the salinity process, EC is mainly due to the use of inorganic fertilizers, organic matter, and pH-improving agents (lime, dolomite). This could be one of the causes of EC fluctuations. When the EC <0.2 mS cm^{-1}, it indicates low salt levels in the soil, which is not good for plants, while the recommended EC range is from 0.2-0.5 mS cm^{-1}. (Mukherjee & Lal. 2014). The results are similar to previous studies, when determining the EC of the orange growing area in the Mekong Delta, the EC fluctuates from 0.08 = 0.58 mS cm^{-1} (Phuong. 2024). This may indicate that the main cause of soil EC change is the use of soil pH improvers, a regular (weekly) activity by farmers to control pH.

CEC in soil samples averaged 28.8 and ranged from 19.0 to 28.8 cmol kg^{-1}. The results obtained were similar to the previous study in Vinh Long, Vietnam, which ranged from 22.6 to 29.5 cmol kg^{-1} (Phuong. 2024).

The average TOC of soil samples was 2.50%, ranging from (1.29-3.34%) in which CL8 had the highest TOC value of 3.34% and the lowest in the CL4 soil sampling area was 1.29%, Table 2. The overall TOC of the study area formed 3 statistically significant distinct areas including (CL4, CL3); (CL1, CL10, CL11, CL6, CL9, CL2, CL5, CL12, CL7) and CL8). The results showed that the TOC value of the study area did not fluctuate too much in the areas. This may be because gardeners paid great attention to using organic matter in soil improvement and organic fertilizers in nutrient supplementation. In general, most soil samples had TOC >1.0%, according to A. Mukherjee and R. L. Lal, TOC was within the average limit (Mukherjee & Lal. 2014). This can be explained by the high cost of organic matter and also by the humid tropical climate with alternating sunshine and rain, which accelerates the process of organic matter decomposition in the soil. Furthermore, the process of cutting water to stimulate off-season flowering also increases the aeration process, which also increases the process of organic matter decomposition. The research results are lower than the previous study in the Mekong Delta orange growing area with an average TOC of 4.2% (Phuong. 2024), this difference may be due to the type of crop and the organic matter addition process in cultivation.

The average available phosphorus of the soil samples was 68.42 mg kg^{-1}, ranging from (25.4-117.1), in which CL5 had the highest available phosphorus content of 117.1 mg kg^{-1} and the lowest of CL3 was 25.4 mg kg^{-1}, Table 2. Most of the soil samples had available phosphorus content >30 mg kg^{-1}. according to Kalu et al., in soils with high P reserves, when the available P in the soil is slightly acidic, adverse environmental impacts may occur on the aquatic environment due to soil erosion. (Kalu et al. 2015). The overall available phosphorus in the study area forms 3 statistically significant distinct zones including (CL8, CL9, CL10); (CL1, CL2, CL3, CL4, CL5, CL6, CL7) and (CL11, CL12). The level of regional division based on available phosphorus is not much. This shows that most gardeners use a lot of phosphorus fertilizer in cultivation to stimulate flowering and branching in the context of acidic soil with low pH and high Al and Fe content.

NH_4^+ content of 12 soil samples had an average content of 39.1 kg^{-1}, ranging from (15.4-74.2), of which CL12 had the highest NH4+ content of 74.2 kg^{-1} and CL10 had the lowest of 15.4 kg^{-1}, Table 2. NH4+ in the entire study area formed 3 statistically significant distinct areas including (CL8, CL9, CL10); (CL1, CL2, CL3, CL4, CL5, CL6, CL7) and (CL11, CL12). The distribution results of NH4+ were similar to that of readily available phosphorus, showing that fertilizer use was regularly focused on by gardeners, so it did not create many zones.

The average clay content of 12 soil samples was 28.8%, ranging from 19.5-36.6, of which (CL 1, CL 4, and CL 9) had the highest clay content of 36.6 and the lowest in the soil sampling area CL 3 and CL 11 was 19.5%, Table 2. The silk content in the soil in the study area averaged 23.3%, the lowest was 12.8 (CL 11), and the highest was 30.8 (CL 1). The soil in the study area is a type of soil with high clay and silk content (loam) suitable for agriculture.

WHC of the soil samples was 71.2%, ranging from 59.5-78.0%, in which CL10 had the highest water holding capacity of 78.0% and the lowest in the CL3 soil sampling area was 59.5%. All samples had a water holding capacity of>50%. According to A. Mukherjee and R. L. Lal, all soil samples could provide adequate water for crops. (Mukherjee & Lal. 2014).

The characteristics of each physicochemical index have been determined and analyzed. However, the correlations of the indexes also need to be considered as a basis for principal component analysis PCA and calculation of SQI value of the soil in the study area.

Table 1 Statistics of results for determining selected soil indices.

	pH		EC mS cm^{-1}		TOC %		Pav mg kg^{-1}		NH_4^+ mg kg^{-1}		BD g cm^{-3}		CEC cmol kg^{-1}		Clay %		Silk %		WHC %	
	Value	SD	Value	SD	Value	SD	Value	SD	Value	SD	Value	SD	Value	SD	Value	SD	Value	SD	Value	SD
CL 1	5.42	0.14	0.05	0.02	2.06	0.02	55.9	12.0	29.4	7.0	0.90	0.02	32.4	0.9	36.6	0.9	30.8	0.8	72.0	1.0
CL 2	4.28	0.07	0.08	0.01	2.64	0.3	85.2	6.3	43.7	3.2	0.92	0.03	30.5	0.8	34.1	0.9	20.5	0.5	67.5	1.5
CL 3	5.14	0.05	0.06	0.02	1.50	0.10	25.4	9.1	35.1	0.8	1.00	0.05	17.7	0.5	19.5	0.5	15.4	0.4	59.5	0.5
CL 4	3.99	0.03	0.11	0.01	1.29	0.01	80.4	2.0	39.4	8.9	0.98	0.00	32.5	0.9	36.6	0.9	28.2	0.7	75.0	1.0
CL 5	4.30	0.01	0.08	0.02	2.98	0.14	117.1	6.3	43.7	0.3	0.99	0.03	30.4	0.8	34.1	0.9	25.6	0.6	69.0	1.0
CL 6	4.92	0.04	0.07	0.01	2.60	0.01	54.2	6.6	33.7	1.9	1.06	0.04	21.6	0.6	24.4	0.6	28.2	0.7	73.0	1.0
CL 7	4.18	0.01	0.11	0.01	3.24	0.04	68.6	1.6	44.8	0.5	0.98	0.02	19.7	0.5	21.9	0.6	20.5	0.5	73.5	0.5
CL 8	4.79	0.01	0.09	0.01	3.34	0.18	85.1	11.2	22.2	1.9	1.04	0.02	21.9	0.6	24.4	0.6	17.9	0.4	71.5	0.5
CL 9	4.65	0.05	0.09	0.01	2.67	0.01	57.3	0.3	23.5	1.3	0.93	0.01	32.6	0.9	36.6	0.9	23.1	0.6	75.5	0.5
CL 10	4.55	0.04	0.10	0.01	2.22	0.02	54.5	0.6	15.4	3.8	0.95	0.02	26.0	0.7	29.3	0.8	28.2	0.7	78.0	1.0
CL 11	4.85	0.02	0.05	0.02	2.38	0.10	74.2	4.1	64.2	2.2	1.05	0.02	17.8	0.5	19.5	0.5	12.8	0.3	65.5	0.5
CL 12	5.02	0.03	0.06	0.01	3.02	0.06	63.0	2.8	74.2	1.3	0.94	0.02	26.0	0.7	29.3	0.8	28.2	0.7	74.5	0.5
Min	3.99		0.05		1.29		25.4		15.4		0.90		19.00		19.5		12.8		59.5	
Max	5.42		0.11		3.34		117.1		74.2		1.06		37.50		36.6		30.8		78.0	
Mean	4.67		0.08		2.50		68.4		39.1		0.98		28.84		28.8		23.3		71.2	

BD: Bulk density, g cm^{-3} ; Pav: available phosphorus, mg $kg^{-1,}$ WHC: The average water holding capacity.

3.2 Correlation analysis results

Table 3 presents the relationships among soil properties, revealing a strong inverse correlation ($p < 0.01$) between soil pH and EC (-0.787), while pH exhibited a perfect positive association with water holding capacity (WHC) at 1.00. Soil EC had a statistically significant negative ($p < 0.01$) relationship with WHC ($r = -0.787$). This may explain why at low pH, some minerals in the soil may be more soluble, releasing ions such as Al^{3+}, Fe^{3+}, Mn^{2+}, etc.

Table 3 Results of correlation analysis of soil indices

	pH	EC	TOC	Pav	NH4+	D	CEC	Clay	Silk	WHC
pH	1									
EC	-0.787**	1								
TOC	-0.088	0.142	1							
Pav	-0.043	-0.392*	0.122	1						
NH4+	0.010	-0.398*	0.156	0.984**	1					
BD	-0.024	-0.021	0.126	0.160	0.094	1				
CEC	-0.252	0.177	-0.140	-0.245	-0.215	-0.659**	1			
Clay	-0.241	0.178	-0.138	-0.247	-0.218	-0.664**	**0.995****	1		
Silk	-0.006	0.128	-0.100	-0.255	-0.173	-0.427**	0.651**	0.655**	1	
WHC	-0.119	0.186	-0.111	-0.256	-0.171	-0.575**	0.773**	**0.779****	0.711**	1
**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed)										

Soil pH negatively correlates with EC, as a decrease in pH raises H^+ ion content and enhances cation leaching (Ca, Mg, Fe, Al, Na, K). Additionally, mineral fertilizer use further lowers pH and increases EC in the acid sulfate soils (Abdel-Fattah et al. 2021). Using pH modifiers like lime and dolomite reduces the mobility of surface ions such as Al and Fe, decreasing EC. Additionally, higher pH enhances organic matter decomposition and soil aeration, resulting in a strong positive correlation between pH and WHC.

Available P showed a significant positive correlation with NH_4^+ ($r = 0.984$, $p < 0.01$), likely due to their combined application in NPK fertilizers. However, the correlation between NH_4^+, available P, and other soil properties was weak and unclear, possibly influenced by factors such as soil texture, TOC, and microbial activity, which can alter these relationships.

Bulk density showed a moderate significant negative correlation with WHC ($r = -0.575$, $p < 0.01$), while clay content had a significant positive correlation with WHC ($r = 0.779$, $p < 0.01$). These changes in soil texture likely result from erosion processes (surface runoff, underground flow, flooding) or the dissolution and accumulation of aluminum and iron. This aligns with Husson et al.'s findings, which highlight the high spatial physical variability of acid sulfate soils (Husson et al. 2000).

3.3 PCA

The analysis of the adequacy and completeness of the data set is based on the Kaiser-Meyer-Olkin (KMO) measure. The KMO value ranges from 0 to 1, with higher values indicating better sample adequacy (see Figure 9.3). Specifically, the KMO value is "0.6 to 0.7 is average, 0.7 to 0.8 is good, 0.8 to 0.9 is excellent, and >0.9 is perfect" (Plonsky. 2015). We have processed and selected the indicators in the data set to have KMO >0.6. The retained quality indicators showed that KMO is 0.63, which is acceptable for PCA analysis, including 8 indicators: pH, EC, Pav, DB, CEC, clay, silk, and WHC. The indicators were removed due to low correlation with many other indicators. Bartlett's test is significant ($p < 0.001$). Table 4 shows that the Sig correlation is close to 0.000 (George & Mallery. 2019; Plonsky. 2015). The KMO and Bartlett test results showed that the data set was suitable for PCA analysis.

Table 4 KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.630
Bartlett's Test of Sphericity	Approx. Chi-Square	289.12
	df	28.000
	Sig.	0.000

The relationship between eigenvalues and PCs is shown in Figure 2. There are three principal factors (PCs) selected with eigenvalues >1 (George & Mallery. 2019). These principal factors cumulatively explain 84.33 % of the variance (Table 5). The results show that the contribution of principal components is sufficient to explain the dataset in the study (Salem & Hussein. 2019). The eigenvalues decrease from PC 1 to PC 3 to 3.98, 1.76, and 1.01, respectively.

The variables with higher loadings are the variables that contribute the most to explaining the meaning of each principal component. The three principal components with the largest percentages of total method variance are 49.78%, 21.95%, and 12.59% of the total method variance, respectively, in Table 5.

With such results, the contribution weight of each PC is calculated according to formula 6 and presented in Table 6, respectively 0.59; 0.26; and 0.15.

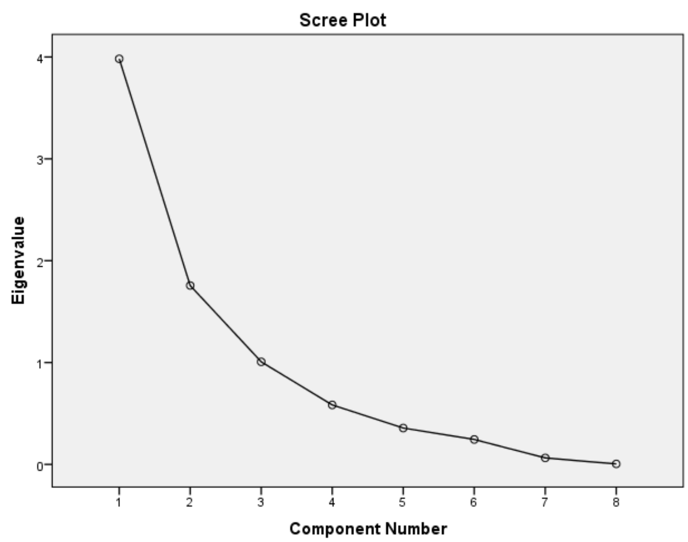

Figure 2 Relationship between eigenvalues and principal components

Table 5 Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.982	49.780	49.780	3.982	49.780	49.780
2	1.756	21.954	71.734	1.756	21.954	71.734
3	1.007	12.592	84.326	1.007	12.592	**84.326**
4	0.584	7.301	91.627			
5	0.357	4.463	96.090			
6	0.245	3.058	99.148			
7	0.064	0.796	99.944			
8	0.005	0.056	100.000			

Extraction Method: Principal Component Analysis.

The principal component analysis results showed that PC1 had the contribution of 7 indices except pH (Table 7. Of these, the high contributing indices were clay content, CEC, and WHC, respectively 0.946; 0.943 and 0.88, Table 7. When considering the correlation between clay and CEC, WHC, they were strongly correlated (0.995 and 0.779, Table 3). Therefore, the clay content index was retained as a representative of PC1. PC1 can be considered a representative of the physical properties of the soil in the study area. The results showed that clay content in soil significantly affects other parameters, such as soil texture, bulk density, water holding capacity, ion exchange capacity, and nutrient retention. A similar explanation is also found in Kome's report. (Kome et al. 2019)

Table 6 Weights of principal components

PC	% of Variance	Weight
1	49.780	0.59
2	21.954	0.26
3	12.592	0.15
	84.326	

Table 7 Component Matrix

Indices	Component		
	1	2	3
pH		0.868	-0.384
EC	0.325	**-0.904**	
Pav	-0.374		**0.882**
DB	-0.719		
CEC	0.943		
Clay	**0.946**		
Silk	0.772		
WHC	0.877		

Extraction Method: Principal Component Analysis.

PC2 has the contribution of 2 indexes (EC and pH) with loading >0.3, Table 7. Of these, EC has the largest loading of 0.90. Since EC and pH have a high negative correlation (r -0.778), only EC is retained. Therefore, EC is retained as a representative for PC2. The results showed that EC reflected the differences in some soil properties, dissolved ion content between soil samples as well as the nutrient absorption capacity of durian trees.

PC3 has a high contribution from the indexes with load >0.3 including available phosphorus, and pH (Table 7. Of these, Pav has the largest load of 0.88 (Table 7), which is representative of PC3. Based on the contributing indexes, this main component can be considered as the factor affecting soil quality. The results showed that the role of valatable P in the development of durian trees is one of the main indicators determining agricultural activities, especially with acid soil.

The expression to calculate the SQI of the soil in the study area is presented in formula (7).

$$SQI = w_{PC1}S_{Clay} + w_{PC2}S_{EC} + w_{PC3}S_{Pav} \quad (6)$$

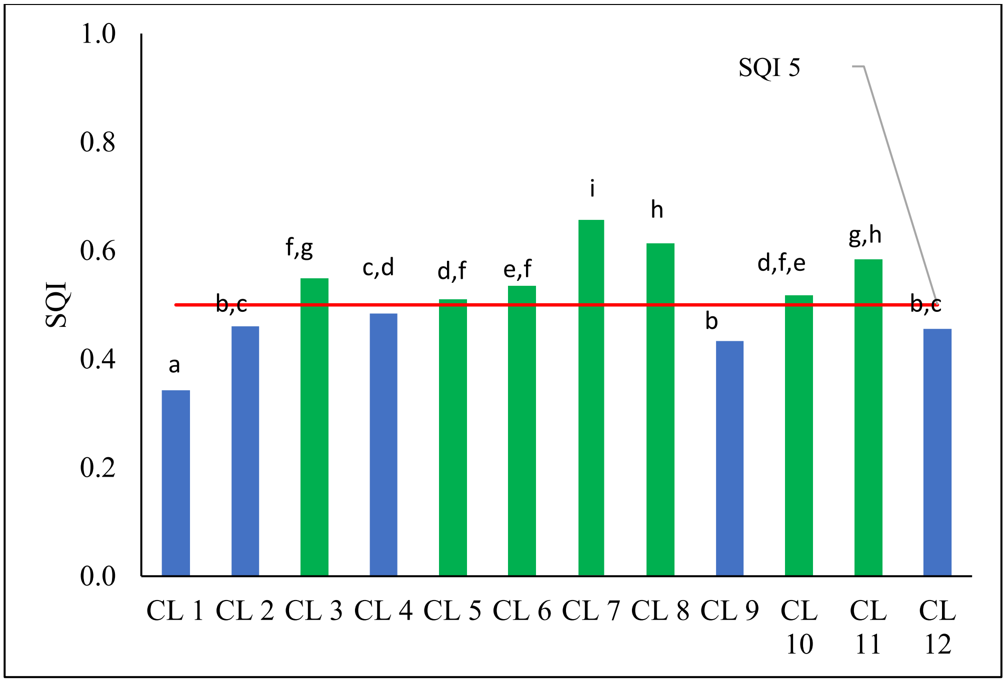

Figure 3 Soil quality index SQI of soil samples, the letters[a,b,c,d,e,f,g,h,i] indicate statistically significant differences in SQI values.

The SQI calculation yielded an average value of 0.51, with the highest (0.66) in sample CL7 and the lowest (0.34) in sample CL1. One-way ANOVA revealed significant SQI differences across the study area, dividing it into 8 distinct sub-locations (Figure 3). According to Damiba et al.'s classification, 83.3% of the 12 locations had moderate SQI (0.4-0.59), 8.3% were good (0.6-0.79), and 8.3% were low (Damiba et al. 2024). The study revealed that soil quality in the area, based on PCA analysis, varied significantly, mostly at a moderate level. Regular monitoring and the use of soil improvers and nutrient supplements are essential to maintain soil health and prevent degradation, which impacts durian productivity. With 83.3% of SQI at moderate levels, government and agricultural organizations should urgently advise farmers on solutions to enhance soil quality and mitigate risks of declining productivity

4 Conclusion

Soil quality was determined by considering 9 indices derived from 12 soil samples collected in the durian growing area of Ben Tre, Vietnam. The selected indices included pH, EC, TOC, Bulk density, CEC, available phosphorus, NH_4^+, clay content, and water holding capacity by creating a minimum data set, and SQI calculation based on PCA was performed. This study showed that most of the soil samples in the study area had a medium soil quality index (SQI). The review of the study data determined that the selected parameters as representatives, such as clay content, EC, and Pav, can be used to determine and monitor soil quality here. PCA can be considered a useful tool for assessing soil quality. Furthermore, the research results also confirmed that regular monitoring of soil quality and the use of interventions in soil quality management to maintain soil quality stability in particular and sustainability in agricultural activities in general are very urgent.

Acknowledgments

The study received assistance from students enrolled in the 16th course of the Institute of Environmental Science, Engineering, and Management. We also extend our gratitude to the reviewers for their valuable feedback, which significantly enhanced the manuscript's quality.

Funding: None

Disclaimer: None

Conflict of Interest: None

ORCID iDs

Nguyen Van Phuong, https://orcid.org/0000-0001-7709-3880

References

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), 1824, DOI: https://doi.org/10.3390/su13041824.

Abdu A., Laekemariam F., Gidago G., & Getaneh L. (2023). Explaining the Soil Quality Using Different Assessment Techniques. *Applied and Environmental Soil Science*, *2023*, DOI: https://doi.org/10.1155/2023/6699154.

Amacher M.C., O'Neill K.P., & Perry C.H. (2007). Soil vital signs: a new soil quality index (SQI) for assessing forest soil health. *Res. Pap. RMRS-RP-65. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 12 p.*, *65*, DOI: https://doi.org/10.2737/RMRS-RP-65.

Amran A., Ariffin M.R., Isa I.M., Ahmed O.H., Herman G., Muhamad S.H., Nor N.A.M., & Khairuddin M.N. (2023). Physicochemical properties of soil cultivated with durian (Durio zibethinus murr.) in Gua Musang, Kelantan. *AGRIVITA Journal of Agricultural Science*, *45*(2), 278-287, DOI: https://doi.org/10.17503/agrivita.v45i2.3601.

Baethgen W., & Alley M. (1989). A manual colorimetric procedure for measuring ammonium nitrogen in soil and plant Kjeldahl digests. *Communications in Soil Science and Plant Analysis*, *20*(9-10), 961-969, DOI: https://doi.org/10.1080/00103628909368129.

Bandyopadhyay S., & Maiti S.K. (2021). Application of statistical and machine learning approach for prediction of soil quality index formulated to evaluate trajectory of ecosystem recovery in coal mine degraded land. *Ecological Engineering*, *170*, 106351, DOI: https://doi.org/10.1016/j.ecoleng.2021.106351.

Bouyoucos G.J. (1962). Hydrometer method improved for making particle size analyses of soils 1. *Agronomy journal*, *54*(5), 464-465, DOI: https://doi.org/10.2134/agronj1962.00021962005400050028x.

Bünemann E.K., Bongiorno G., Bai Z., Creamer R.E., De Deyn G., de Goede R., Fleskens L., Geissen V., Kuyper T.W., Mäder P., Pulleman M., Sukkel W., van Groenigen J.W., & Brussaard L. (2018). Soil quality – A critical review. *Soil Biology and Biochemistry*, *120*, 105-125, DOI: https://doi.org/10.1016/j.soilbio.2018.01.030.

Damiba W.A.F., Gathenya J.M., Raude J.M., & Home P.G. (2024). Soil quality index (SQI) for evaluating the sustainability status of Kakia-Esamburmbur catchment under three different land use types in Narok County, Kenya. *Heliyon*, *10*(5), DOI: https://doi.org/10.1016/j.heliyon.2024.e25611.

George D., & Mallery P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference*. Routledge. https://doi.org/https://doi.org/10.4324/9780429056765

Giao N.T., Nhien H.T.H., Dan T.H., & Ni D.V. (2023). Identifying Key Parameters Influencing Soil Quality at Various Depths in Tram Chim National Park, Dong Thap Province, Vietnam. *Journal of Ecological Engineering*, *24*(2), 81-91, DOI: https://doi.org/10.12911/22998993/156757.

Husson O., Verburg P., Phung M.T., & Van Mensvoort M. (2000). Spatial variability of acid sulphate soils in the Plain of Reeds, Mekong delta, Vietnam. *Geoderma*, *97*(1-2), 1-19, DOI: https://doi.org/10.1016/S0016-7061(00)00016-1.

Ivezić V., Singh B.R., Gvozdić V., & Lončarić Z. (2015). Trace Metal Availability and Soil Quality Index Relationships under Different Land Uses. *Soil Science Society of America Journal*, DOI: https://doi.org/10.2136/sssaj2015.03.0125.

Kalu S., Koirala M., Khadka U.R., & Anup K. (2015). Soil quality assessment for different land use in the Panchase area of western Nepal. *International Journal of Environmental Protection*, *5*(1), 38-43, DOI: https://doi.org/10.5963/IJEP0501006.

Kome G.K., Enang R.K., Tabi F.O., & Yerima B.P.K. (2019). Influence of clay minerals on some soil fertility attributes: a review. *Open Journal of Soil Science*, *9*(9), 155-188, DOI: https://doi.org/10.4236/ojss.2019.99010.

Mukherjee A., & Lal R.L. (2014). Comparison of Soil Quality Index Using Three Methods. *PLoS ONE*, *9*(8), DOI: https://doi.org/10.1371/journal.pone.0105981.

Nguyen B.T.N., Kansal M.L., & Nguyen H.-H. (2024). Soil Quality Assessment Towards its Sustainable Management in Thai Binh Mangrove, Vietnam. *Water, Air, & Soil Pollution*, *235*(8), 479, DOI: https://doi.org/10.1007/s11270-024-07259-2.

Pham T.G., Nguyen H.T., & Kappas M. (2018). Assessment of soil quality indicators under different agricultural land uses and topographic aspects in Central Vietnam. *International Soil and Water Conservation Research*, *6*(4), 280-288.

Phuong N.V. (2024). EVALUATION OF QUALITY OF ORANGE GROWTH LAND IN TRA ON DISTRICT, VINH LONG PROVINCE. *Journal of Science and Technology-IUH*, *68*(02), DOI: https://doi.org/10.46242/jstiuh.v68i02.5086.

Plonsky L. (2015). Advancing quantitative methods in second language research. In: Routledge.

Rajput R.S., Patley S., Choudhary B., & Gupta S. (2023). Assessment and Comparison of Soil Quality Using Principal Components Analysis (PCA) & Expert Opinion (EO) Methods in different Rice-based Cropping Systems in Alfisol. *International Journal of Environment and Climate Change*, *13*(12), 869-887, DOI: https://doi.org/10.9734/ijecc/2023/v13i123750.

Rangel-Peraza J.G., Padilla-Gasca E., López-Corrales R., Medina J.R., Bustos-Terrones Y., Amabilis-Sosa L.E., Rodríguez-Mata A.E., & Osuna-Enciso T. (2017). Robust Soil Quality Index for Tropical Soils Influenced by Agricultural Activities. *Journal of Agricultural Chemistry and Environment*, *6*, 199-221.

Salem N., & Hussein S. (2019). Data dimensional reduction and principal components analysis. *Procedia Computer Science*, *163*, 292-299, DOI: https://doi.org/10.1016/j.procs.2019.12.111.

Van Binh N., Duc T.T., & Hoa H.T.T. (2013). Assessment of Soil Quality Parameters under Rice and Maize Cropping Systems in Hilly Region of Central Vietnam. *Indian Journal of Agricultural Research*, *1*, 6, DOI: https://doi.org/10.18805/IJARe.AF-883.