Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

p-ISSN: 0972-6268 (Print copies up to 2016)
e-ISSN: 2395-3454

Vol. 25 | No. 3 | Article ID D1893 | 2026

Original Research Paper

https://doi.org/10.46488/NEPT.2026.v25i03.D1893

Open Access Journal

Application of Principal Component Analysis (PCA) and Interpolated Spatial Distribution Maps in Soil Quality Control of Dragon Fruit Farms in Tay Ninh Province, Vietnam

Nguyen Van Phuong†, Tong Hai Minh and Le Hong Thia
Institute of Environmental Science, Engineering & Management, Industrial University of Ho Chi Minh City, 70000, Vietnam
†Corresponding author: Nguyen Van Phuong; nguyenvanphuong@iuh.edu.vn

Abbreviation: Nat. Env. & Poll. Technol.
Website: www.neptjournal.com

Received: 20-12-2025
Revised: 17-02-2026
Accepted: 26-02-2026

Key Words:

Dragon fruit
Soil quality index (SQI)
Principal Component Analysis (PCA)
Minimum dataset (MDS)
Soil quality control

Citation for the Paper:

Phuong, N.V., Minh, T.H. and Thia, L.H., 2026. Application of Principal Component Analysis (PCA) and interpolated spatial distribution maps in soil quality control of dragon fruit farms in Tay Ninh Province, Vietnam. *Nature Environment and Pollution Technology*, 25(3), D1893.

Note: From 2025, the journal has adopted the use of Article IDs in citations instead of traditional consecutive page numbers. Each article is now given individual page ranges starting from page 1.

ABSTRACT

Accurate and timely assessment of soil quality is essential for sustainable farming and efficient agricultural management. The Soil Quality Index (SQI) is heavily influenced by fluctuations in physical and chemical soil indicators. Twelve composite soil samples were collected from the dragon fruit cultivation area in Tay Ninh, Vietnam. The dataset included eleven soil parameters: pH, electrical conductivity (EC), total organic carbon (TOC), cation exchange capacity (CEC), available phosphorus (P_av), ammonium (NH4+), bulk density, particle density, clay, silt, and sand. Principal Component Analysis (PCA) was employed to identify the Minimum Dataset (MDS) and determine parameter weights. These were combined with land index scores to estimate the SQI. Additionally, a spatial distribution map of SQI was created using the Inverse Distance Weighting (IDW) method in ArcGIS 10.8. Results showed that the MDS is influenced by four principal components, which explain 89.02% of the total data variance. Key parameters include CEC, P_av, EC, and pH, with weights of 0.37, 0.26, 0.22, and 0.15, respectively. The average soil quality score was 41.6%, with 58.3% classified as degraded. The study demonstrated that combining PCA and GIS provides a comprehensive and intuitive approach to SQI evaluation. Furthermore, developing agronomic maps based on large sample sizes supports sustainable agricultural management, as evidenced by these research findings.

1. INTRODUCTION

Soil quality determines the ability of soil to support plant health, maintain productivity, and sustain farmers' livelihoods and economic stability (Rajput et al. 2023). Comprehensive monitoring of soil health and a thorough understanding of soil indicators and their essential functions are critical for developing effective solutions and policies aimed at sustainable land management (Țopa et al. 2025).

Various approaches have been established to quantify and monitor soil quality based on the calculation of indices, including the Soil Quality Index (SQI), Soil Health Index (SHI), and Soil Sustainability Index (SSI), which are used as decision-support tools (Damiba et al. 2024). However, none of these indices has been widely accepted for soil quality assessment, possibly because of the complexity and heterogeneity of soil systems. Three approaches, namely the Total Data Set (TDS), Expert Data Set (EDS), and Minimum Data Set (MDS), are commonly used for soil quality assessment (Damiba et al. 2024). In the TDS approach, the SQI is calculated as the average score of all selected indicators. In the EDS approach, indicator weights are assigned based on expert knowledge, although sufficient information is often unavailable for all soil indicators. In contrast, the MDS approach is based on Principal Component Analysis (PCA). Compared with the TDS and EDS approaches, the MDS requires less intensive field and laboratory work, thereby reducing both time and cost (Damiba et al. 2024).

In addition, many studies have focused on individual physical and chemical indicators, although these may provide an incomplete assessment because many soil properties result from interactions among several processes. In some studies, biological indicators have been ignored or excluded (Damiba et al. 2024, Ţopa et al. 2025). Priority soil indicators are generally selected based on the following criteria: ease of measurement, rapid analysis, low cost, applicability under field conditions, and sensitivity to changes in management practices (Ţopa et al. 2025). Common soil indicators used in soil quality assessment include physical properties such as bulk density, particle density (specific gravity), water retention, and the percentages of clay, sand, and silt, as well as chemical properties including pH, electrical conductivity (EC), organic matter, cation exchange capacity (CEC), NH_4^+, available phosphorus (P_av), copper, and zinc (Abdel-Fattah et al. 2021, Hanif et al. 2025, Rajput et al. 2023, Van 2025).

The MDS-based SQI calculation method is currently the most widely used because of its flexibility, quantitative nature, and ease of application. Statistical techniques such as PCA and Pearson correlation analysis are commonly employed to develop the MDS (Rajput et al. 2023, Ţopa et al. 2025, Van 2025). This method provides not only an SQI value but also a minimum set of soil parameters for monitoring soil health (Rajput et al. 2023). PCA is a valuable statistical tool for data reduction and improving model efficiency. It removes correlated variables, enhances the statistical significance of selected indicators, and minimizes multicollinearity. By transforming high-dimensional datasets into a smaller number of principal components, PCA reduces the number of variables included in the model, thereby improving computational efficiency and facilitating data visualization and interpretation (Abdel-Fattah et al. 2021). The Inverse Distance Weighting (IDW) interpolation method in ArcGIS 10.8 is widely used to estimate SQI values at unsampled locations based on surrounding sample points and to map the spatial distribution of soil quality (Abdel-Fattah et al. 2021, Kahsay et al. 2025).

Tay Ninh Province is the second-largest dragon fruit-producing province in Vietnam, with more than 7,000 ha under cultivation (Mai et al. 2025). The cultivated area is characterized by acid sulfate soils within the Vam Co River basin. However, soil quality management in this dragon fruit-growing region still lacks effective solutions to ensure sustainable production, including precision farming and modern management practices. Therefore, this study combines PCA and Geographic Information Systems (GIS) to identify key soil quality indicators and simulate the spatial variation of SQI in dragon fruit farms in Tay Ninh Province, Vietnam, using a dataset comprising 11 soil parameters: pH,

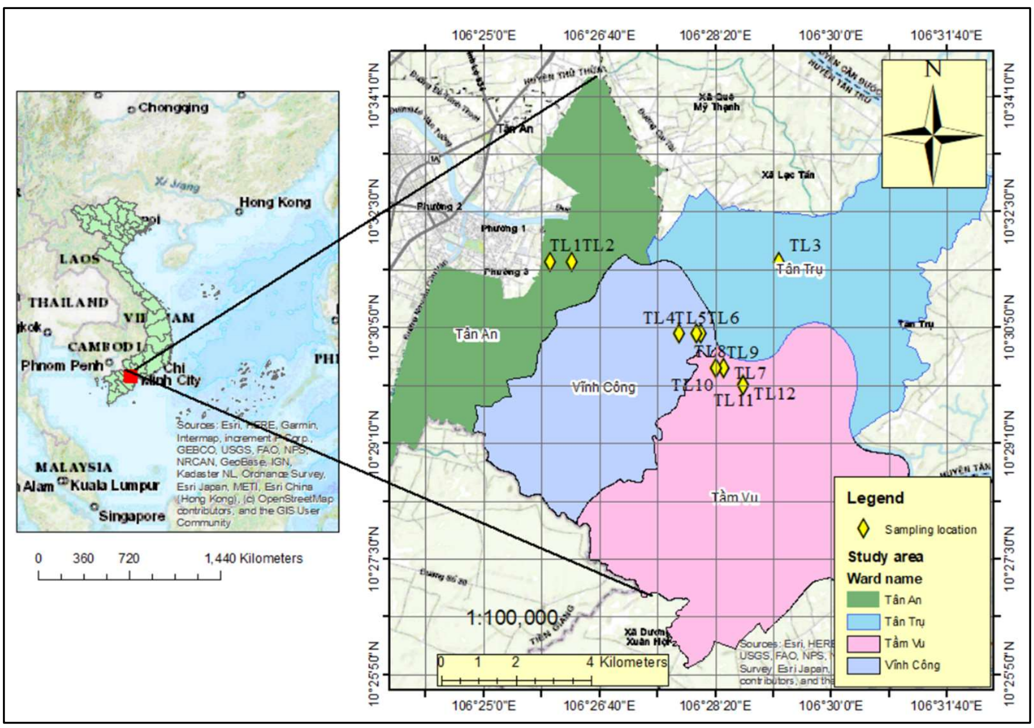


Fig. 1: Map of sampling locations in dragon fruit farms in Tay Ninh, Vietnam (TL1 to TL12 are 12 composite samples collected in the dragon fruit farms, Tay Ninh, Vietnam).

EC, TOC, P_av, NH_4^+, bulk density (d), particle density (D), clay content, silt content, sand content, and CEC.

2. MATERIALS AND METHODS

2.1. Field Sampling

In January 2024, sixty soil samples were collected from dragon fruit farms in Tay Ninh province to form twelve composite samples (see Fig. 1). Each selected area, roughly 4 km² (400 hectares), displayed the greatest homogeneity in dragon fruit tree age and color. Within each zone, a sampling site of at least 1.5 hectares was designated. Soil samples were extracted at a depth of 0-30 cm, the layer most indicative of soil quality dynamics. The sampling employed a composite method: in each area exceeding 1.5 hectares, samples were taken from the four corners and the center along the diagonal. The collected soil was naturally dried, ground finely, and sieved through a 2 mm mesh before analyzing eleven soil parameters: pH, electrical conductivity (EC), total organic carbon (TOC), cation exchange capacity (CEC), available phosphorus (P_av), ammonium (NH4+), bulk density, particle density, clay, silt, and sand.

2.2. Analytical Methods

The soil properties were determined using standardized methods. Soil particle density (d) was determined according to TCVN 6863:2001, pH and electrical conductivity (EC) were measured following ISO 10390:1993, bulk density (D) was determined according to TCVN 8305:2009, total organic carbon (TOC) was analysed using the Walkley–Black method, available phosphorus (P_av) was determined using the Olsen method, ammonium (NH_4^+) was determined following the method of Baethgen & Alley (1989), soil texture was analyzed using the hydrometer method (George 1962), and cation exchange capacity (CEC) was determined using the ammonium acetate method according to TCVN 8568:2010. Analytical-grade reagents, including H_2SO_4, $K_2Cr_2O_7$, $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$, Na_3PO_4, $SnCl_2$, and $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ (Merck, Germany, and China), were used for all laboratory analyses.

2.3. Soil Quality Assessment

Principal Component Analysis (PCA) is a multivariate statistical technique employed for data reduction by simplifying data structure through the extraction of principal components (PCs) and analyzing the orthogonal relationships among variables. Following PCA, the Soil Quality Index (SQI) is determined through a series of steps, which include standardizing variables, constructing a correlation matrix, identifying principal components and their explained variance, calculating indicator weights, analyzing factor loadings, and ultimately computing the SQI (Aisah et al. 2025, Damiba et al. 2024). Since soil indicators are measured in different units, they are standardized to a 0–1 scale to facilitate comprehensive soil quality assessment and to ensure all indicators contribute equally (Damiba et al. 2024). Based on reference values and expert judgment, three common scoring functions, linear or non-linear, are used: "more is better," "optimal range," and "less is better" (Bünemann et al. 2018). In this study, the "more is better" scoring function was applied to relevant soil indicators and is given by Equation (1) (Bandyopadhyay & Maiti 2021).

$$S_i = \frac{X}{X_{max}} \quad ...(1)$$

Where X, Xmax, and Xmin are the measured value, the maximum value, and the minimum value. Si is the score of indicator i. Non-linear functions, including pH, EC, clay, silt, and sand, and linear (NH4+, P_av, TOC, CEC, the more the better), the less the better (bulk density, particle density), are presented in detail in Table 1 (Damiba et al. 2024, Kahsay et al. 2025)

For the soil indices: "less is better" is described in formula (Eq. 2):

$$S_i = \frac{X_{min}}{X} \quad ...(2)$$

For the non-linear scoring function, the soil parameters are converted according to the S-shaped curve equation; formula (Eq. 3) is as follows:

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

Where Xi is the measured value, Ximean is the average value, and Xopt is the optimal value (Table 1). Slope b = -2.5 (more is better), b = +2.5 (less is better) và Xo (Damiba et al. 2024, Kahsay et al. 2025)

A Pearson correlation matrix was generated to evaluate the relationships among the variables. Correlation coefficients

Table 1: Components of the scoring function.

Indices	Scoring curve	Xmax	Xmin	Xopt
Sand [%]	Optimum	90.00	6.00	36
Silt [%]	Optimum	50.00	4.00	30
Clay [%]	More is better	84.00		
D [g.cm^{-3}]	Less is better	66.00	0.89	
pH [H_2O]	Optimum	9.62	5.04	7.00
EC [dS m^{-1}]	Less is better	2	0.001	1
TOC [%]	More is better	5.96		
P_av [mg.kg^{-1}]	More is better	58.89		
CEC [cmoL.kg^{-1}]	More is better	61.98		

range from −1 to +1, where the magnitude indicates the strength of the relationship and the sign denotes its positive or negative nature. These correlations serve as a foundation for the subsequent Principal Component Analysis (PCA). PCA reduces the dataset's dimensionality by transforming the original variables into a set of uncorrelated principal components (PCs) that explain most of the original variance. Variables with the highest loadings on these PCs are considered the most influential parameters affecting agricultural land quality (Rangel-Peraza et al. 2017). Communalities represent the percentage of variance each variable explains in the PCs. PCs with eigenvalues greater than 1 and accounting for at least 10% of the data's variation are retained. From these, a Minimum Data Set (MDS) is derived by selecting variables with significant factor loadings. These influential variables have either the greatest weights on their respective PCs or factor loadings within 10% of the highest loading (Abdu et al. 2023, Damiba et al. 2024). When multiple indicators are grouped within the same PC, the correlation matrix is used to assess inter-variable relationships ($p < 0.05$). If the correlation is strong ($r > 0.5$), the indicator with the highest factor loading is chosen. If the correlation is weak ($r < 0.5$), all indicators are retained (Damiba et al. 2024).

2.4. SQI Calculation and Classification

Calculate the Soil Quality Index Plus (SQI_Plus), formula 4 (Eq. 4) (Damiba et al. 2024)

$$SQI_{Plus} = \frac{\sum_{i=1}^{n} S_i}{n} \quad ...(4)$$

SQI is calculated according to formula (Eq. 5):

$$SQI = \sum_{i=1}^{n} W_i . S_i \quad ...(5)$$

In which: n is the total number of related indicators, Wi is the weight of PCi, and Wi is calculated according to formula (Eq. 6) (Damiba et al. 2024). Where PCi is the rotated load of principal component i and SPC is the total rotational load.

$$W_i = \frac{PC_i}{\sum PC} \quad ...(6)$$

Classification for SQI levels was assessed according to the scale used by Li et al, where SQI>0.8 (very good), 0.7-0.8 (good), 0.6-0.7 (moderate), 0.4-0.6 (degraded), and <0.4 (poor) (Li et al. 2024)

2.5. Data Processing

To ensure reliability, each analysis was performed in triplicate, and the mean, standard deviation, and range were calculated using Excel. Data suitability for PCA was assessed using Bartlett's test and the Kaiser-Meyer-Olkin (KMO) measure, where values above 0.5 are considered acceptable, and Pearson correlation analysis was used to examine the relationships among soil indices using SPSS 23.0 software (Abdel-Fattah et al. 2021, Plonsky 2015). The spatial distribution map of the SQI was generated using ArcGIS 10.8 and the inverse distance weighting (IDW) interpolation method.

3. RESULTS AND DISCUSSION

3.1. Some Soil Properties in the Study Area

The results of the determination of 11 soil physicochemical properties in the dragon fruit-growing area of Tay Ninh, Vietnam, are presented in Table 2. Specifically, the pH values ranged from 6.07 to 7.21, with an average value of 6.48, indicating that the soil pH of the study area is highly suitable for plant growth (Mukherjee & Lal 2014). This suggests that farmers are aware of the importance of controlling soil pH, although previous studies have shown that the soils in this area are generally characterized as acid sulfate soils (pH 4.0–5.4) (Mai et al. 2025, Van 2025).

The results showed that the study area was characterized by low soil EC, with an average value of 0.31 mS cm^{-1}, ranging from 0.16 to 0.45 mS cm^{-1}. According to Mukherjee and Lal, EC values of 0.2–0.5 mS cm^{-1} are considered most suitable for crop cultivation (Mukherjee & Lal 2014). The average TOC content was 3.28%, ranging from 1.80 to 5.04%, indicating a medium to high organic matter content (Mukherjee & Lal 2014). However, because dragon fruit is a perennial crop, insufficient application of organic amendments by farmers may lead to a decline in soil organic matter over time.

The CEC of the study area varied widely, ranging from 11.1 to 33.7 cmoL.kg^{-1} soil, with an average value of 22.5 cmoL.kg^{-1}. These results were similar to those reported for soils in Ben Tre, Vietnam, where CEC ranged from 19.0 to 28.8 cmoL.kg^{-1} (Van 2025). The average available P content was 37 mg.kg^{-1}, ranging from 12 to 66 mg.kg^{-1}, which was lower than that reported for durian-growing soils, where available P ranged from 25.4 to 117.1 mg.kg^{-1} (Van 2025). This difference may be attributed to the different phosphorus requirements of each crop, as well as differences in nutrient management practices associated with production costs.

The average NH_4^+ content was 114 mg.kg^{-1}, ranging from 44 to 148 mg.kg^{-1}, indicating that the NH_4^+ content was similarly high compared with that of neighboring soils reported in a previous study (Van 2025). The lowest NH_4^+ concentration was observed in soil sample TL6, where the crop had just been harvested, whereas the highest

Table 2: Soil physical and chemical indices in the study area.

	pH	EC mS cm^{-1}	TOC %	CEC cmoL.kg^{-1}	P_av mg.kg^{-1}	NH_4^+ mg.kg^{-1}	d g.cm^{-3}	D g.cm^{-3}	Clay %	Silt %	Sand %
TL1	6.17	0.35	3.62	11.1	57	125	2.35	1.06	13	28	60
SD	0.01	0.03	0.14	0.5	8	1	0.13	0.01	1	1	2
TL2	7.15	0.28	3.16	15.7	20	137	2.67	0.91	18	23	60
SD	0.06	0.03	0.16	0.7	1	2	0.07	0.01	1	1	3
TL3	6.92	0.29	3.48	22.5	28	104	2.55	0.98	25	18	58
SD	0.04	0.00	0.24	0.8	6	6	0.02	0.01	1	1	3
TL4	6.77	0.32	4.88	24.7	57	148	2.33	1.03	28	18	55
SD	0.02	0.01	0.16	0.9	2	9	0.16	0.01	1	1	2
TL5	6.60	0.38	4.00	22.5	51	143	2.21	1.01	25	18	58
SD	0.00	0.01	0.28	0.8	3	3	0.09	0.00	1	1	2
TL6	6.46	0.16	3.02	20.4	53	44	2.73	1.15	23	13	65
SD	0.06	0.02	0.02	0.0	2	7	0.02	0.03	1	1	3
TL7	6.32	0.22	2.78	33.7	23	63	2.43	1.10	38	15	48
SD	0.05	0.01	0.10	0.8	1	2	0.04	0.01	1	1	2
TL8	6.25	0.31	3.27	29.3	42	90	2.08	1.01	33	13	55
SD	0.01	0.01	0.13	0.8	6	9	0.01	0.08	1	1	2
TL9	6.45	0.44	3.18	22.5	37	121	2.57	0.92	25	18	58
SD	0.04	0.02	0.18	0.1	2	4	0.02	0.05	1	1	3
TL10	6.22	0.35	3.18	24.7	36	135	2.22	1.05	28	18	55
SD	0.09	0.03	0.18	0.2	1	5	0.11	0.03	1	1	2
TL11	6.31	0.33	2.90	22.5	16	116	2.26	0.91	25	18	58
SD	0.04	0.02	0.10	0.3	1	4	0.22	0.01	1	1	3
TL12	6.12	0.36	1.83	20.4	20	138	2.72	0.93	23	13	65
SD	0.04	0.01	0.03	0.5	8	2	0.01	0.01	1	1	3
Min	6.07	0.16	1.80	11.1	16	44	2.04	0.87	13	13	48
Max	7.21	0.45	5.04	33.7	57	148	2.75	1.19	38	28	65
Average	6.48	0.31	3.28	22.5	37	114	2.43	1.01	25	17	58

SD: standard deviation.

concentration was recorded in sample TL4 during the fruit development stage.

Soil particle density ranged from 2.04 to 2.75 g.cm^{-3}, with an average value of 2.43 g.cm^{-3}. Soil bulk density ranged from 0.87 to 1.19 g.cm^{-3}, corresponding to light- to medium-textured soils, which are common in the Mekong Delta region of Vietnam (Van 2025). The clay, silt, and sand contents ranged from 13 to 38%, 13 to 28%, and 48 to 65%, respectively. The soil texture was classified as loam, which is suitable for cultivation.

3.2. Pearson Correlation Analysis

The correlation coefficients among the soil parameters are presented in Table 3. Soil pH showed low correlations with all other soil parameters. This indicates that the application of alkaline materials, such as lime (CaO and $CaCO_3$), for pH control and regular soil management has little impact on other soil parameters. This finding is consistent with the soil pH results (Table 2), which showed only minor variations in pH.

Soil EC showed a significant positive correlation ($p < 0.05$) with NH_4^+ ($r = 0.80$). The results indicate that the application of nitrogen fertilizers contributed to an increase in the concentration of soluble salts in the soil. Soil TOC was strongly correlated with available P ($r = 0.65$, $p < 0.01$), suggesting that the mineralization of organic matter contributed to an increase in available P through desorption. The significant variation in TOC was attributed to differences

Table 3: Correlation coefficients of soil parameters.

	pH	EC	TOC	P_av	NH_4^+	d	D	% clay	% silt	% sand	CEC
pH	1										
EC	-0.165	1									
TOC	0.454**	0.150	1								
P_av	-0.036	0.002	0.654**	1							
NH_4^+	0.176	0.799**	0.308	0.012	1						
D	-0.259	0.286	0.373*	0.254	0.171	1					
D	0.244	0.580**	-0.210	-0.559**	0.563**	-0.117	1				
% clay	-0.162	-0.212	-0.039	-0.202	-0.376*	0.344*	-0.229	1			
% silt	-0.073	-0.125	-0.030	-0.225	-0.275	0.283	-0.087	0.945**	1		
% sand	-0.323	-0.343*	-0.417*	-0.131	-0.507**	-0.052	-0.197	0.621**	0.698**	1	
CEC	-0.163	-0.220	-0.055	-0.210	-0.386*	0.331*	-0.224	0.999**	0.949**	0.639**	1

**. Correlation is significant at the 0.01 level (2-tailed); *. Correlation is significant at the 0.05 level (2-tailed)

in organic matter management strategies in crop production (Topa et al. 2025).

CEC showed significant positive correlations ($p < 0.01$) with clay ($r = 1.00$), silt ($r = 0.95$), and sand ($r = 0.64$). The results indicate that soil texture strongly influenced soil CEC. Similar findings were reported by Kawaguchi and Kyuma, who suggested that clay minerals explain more than 80% of the variation in CEC, whereas the contribution of organic matter is not statistically significant in paddy soils of tropical Asia (Kawaguchi & Kyuma 1975).

3.3. Principal Component Analysis

Table 4 presents the Bartlett and KMO test results of the total data set (TDS). Bartlett test shows a significance level <0.0001 and an observed chi-square value (607.619), p-value = 0.000 < 0.05, which indicates that the variables are significantly correlated with each other, and the data are suitable for PCA analysis. In addition, the KMO value is also above 0.538 (>0.50), confirming that the sample size is reliable enough for factor analysis. These results confirm that the data set meets the conditions for performing PCA. (Abdel-Fattah et al. 2021).

The relationship between eigenvalues and principal components (PCs) is illustrated in Fig. 2. The analysis identified four PCs with eigenvalues greater than 1, specifically, 4.12, 2.19, 2.14, and 1.35 that together explain over 89.02% of the variance in the original data (Table 5). The contribution levels of these PCs are 33.2%, 22.7%, 19.9%, and 13.2%, respectively (Table 5).

Table 4: Bartlett and KMO test.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.538
Bartlett's Test of Sphericity	Approx. Chi-Square	607.619
	df	55
	Sig.	0.000

W_PC1, W_PC2, W_PC3, and W_PC4 are the weights of the respective PCs, calculated according to Eq. 6.

PC1 explained 33.2% of the total variance. Parameters with absolute loading values > 0.6 included CEC, clay, silt, sand, and NH_4^+, with loadings of 0.916, 0.909, 0.864, 0.817, and −0.659, respectively, indicating that these parameters made major contributions to PC1 (Table 6). Among them, CEC had the highest loading value, whereas the loading values of clay and silt were within 10% of the CEC loading value and were therefore considered for selection (Abdu et al. 2023, Damiba et al. 2024). However, these two indices were significantly correlated ($P < 0.01$) with CEC, with correlation coefficients (r) of 1.00 and 0.95, respectively (Table 3), and were therefore excluded to avoid multicollinearity (Damiba et al. 2024). Consequently, only CEC was retained as the representative variable for PC1. The results showed that PC1 represents the soil texture component, characterizing the weathering process of the parent material and the effects of soil erosion, which may be caused by rainfall or irrigation.

PC2 explained 22.7% of the total variance. Parameters with loading values > 0.6 included P_av, TOC, and bulk density (D), with loadings of 0.861, 0.830, and 0.643, respectively. Among these, P_av had the highest loading value, while the loading value of TOC was within 10% of that of P_av; therefore, these two parameters were considered for PC2. However, a strong correlation was observed between P_av and TOC ($r = 0.654$, $p < 0.01$, Table 3), and only P_av was retained as the representative variable for PC2 (Damiba et al. 2024). The results indicated that available P was mainly

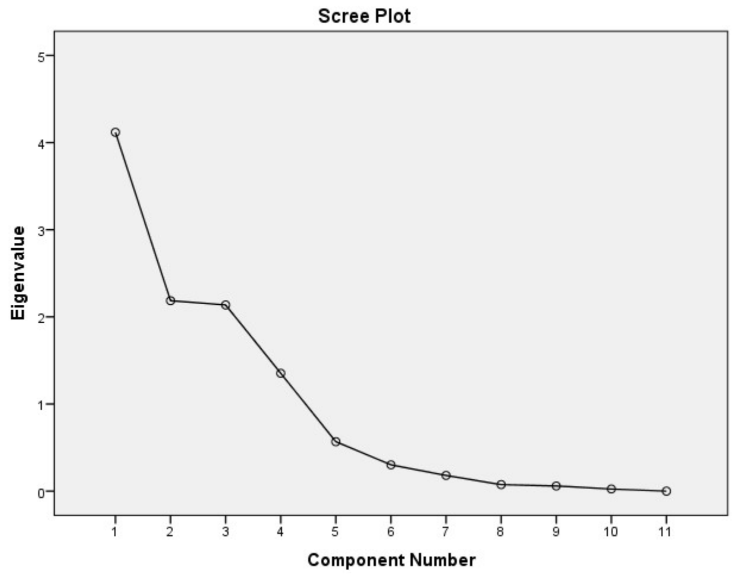


Fig. 2: Eigenvalues by Principal Components.

Table 5: PCs explain total variance.

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.118	37.439	37.439	3.655	33.223	33.223
2	2.185	19.867	57.306	2.497	22.703	55.926
3	2.136	19.420	76.725	2.184	19.850	75.776
4	1.352	12.295	89.020	1.457	13.244	89.020
5	0.568	5.159	94.179			
6	0.302	2.744	96.924	W_PC1	0.37	
7	0.180	1.637	98.560	W_PC2	0.26	
8	0.075	0.681	99.241	W_PC3	0.22	
9	0.059	0.538	99.780	W_PC4	0.15	
10	0.024	0.220	100.000			
11	2.588E-05	.000	100.000			

Extraction Method: Principal Component Analysis.

Table 6: Component matrix.

	Component			
	1	2	3	4
pH	-0.294			0.932
EC	-0.474		**0.762**	-0.294
TOC	-0.268	0.830	0.241	0.369
P_av	-0.188	**0.861**	-0.181	-0.197
d	0.174	0.526	0.588	-0.294
% silt	0.864		0.416	0.210
CEC	0.916		0.334	0.135
% sand	0.817	-0.193		-0.116
D	-0.393	-0.643	0.535	0.204
% clay	0.909		0.343	0.135
NH_4^+	-0.659		0.652	

Extraction Method: Principal Component Analysis.

derived from fertilizer application. However, because the study soils are acidic and have a high phosphorus fixation capacity, the presence of organic matter may contribute to increasing the availability of phosphorus. Therefore, PC2 can be considered representative of the combined effects of phosphorus and organic matter.

PC3 accounted for 19.9% of the total variance. Parameters with loading values > 0.6 included EC and NH_4^+, with loadings of 0.762 and 0.652, respectively. Among these, EC had the highest loading value, while the loading value of NH_4^+ was within 10% of that of EC; therefore, these two parameters were considered for PC3. However, EC and NH_4^+ were strongly correlated ($r = 0.799$, $p < 0.01$, Table 3), and only EC was retained as the representative variable for PC3 (Damiba et al. 2024). PC3 is considered to represent soil mineralization and ammonium nitrogen application.

PC4 explained 13.2% of the total variance. Soil pH was the only parameter with a loading value of 0.932 (Table 6) and was therefore retained as the representative variable for PC4 (Damiba et al. 2024). PC4 is considered to represent the application of soil amendments, mainly lime, for soil pH management, according to information obtained from farmers. PC1, PC2, PC3, and PC4, represented by CEC, P_av, EC, and pH, respectively, were selected as the Minimum Data Set (MDS), with a cumulative contribution of 89.02%.

The weighted values (W_PC), representing the contribution of each principal component, were calculated using Eq. 6. The resulting weights were 0.37, 0.26, 0.22, and 0.15 for PC1, PC2, PC3, and PC4, respectively (Table 6). The general formula for calculating the SQI (Eq. 5) for the sampling locations in the study area is presented in Eq. 7:

$$\boldsymbol{SQI = \sum_{i=1}^{n} W_i . S_i = 0.37 * S_{CEC} + 0.26 * S_{P_{av}} + 0.22 * S_{EC} + 0.15 * S_{pH}} \quad ...(7)$$

Calculate Si scores for soil indices: Table 7 presents the results of Si scoring using linear, non-linear, and optimized scoring functions (Eq. 1, Eq. 2, and Eq. 3). The findings indicate that, except for soil particle density and bulk density Si scores, both exceeding 0.5 with ranges of 0.75 to 0.98 and 0.76 to 0.96 respectively, most other soil indices exhibit significant fluctuations and signs of degradation, with scores below 0.5. The pH scores, ranging from 0.46 to 0.56, are concerning and warrant close monitoring. Notably, indices with considerable variations include EC (0.16 to 0.70), P_av (0.25 to 0.87), NH4+ (0.28 to 0.94), CEC (0.31 to 0.87), and clay (0.33 to 1.00). These variations may serve as indicators of soil index degradation.

This study estimated the SQI of 12 Dragon Fruit cultivation areas using two calculation methods: the Average Addition SQI_Plus (derived from Eq. 4) and the PCA analysis SQI_PCA (from Eq. 7), as shown in Fig. 3. The SQI_Plus has an average of 0.65, ranging from 0.57 to 0.75, with the lowest in sample TL 2 and the highest in TL 4. The SQI_PCA ranges from 0.43 to 0.69, with an average of 0.57, also lowest in TL 2 and highest in TL 4 (Fig. 3). The two methods show highly consistent variation in SQI values, despite differences in their specific results. Both methods support the same overall assessment of SQI. Notably, SQI_PCA values are consistently lower than SQI_Plus, likely

Table 7: Si calculation results based on soil index scoring functions.

	pH	EC	TOC	P_av	NH_4^+	d	D	clay	silt
TL 1	0.47	0.56	0.72	0.87	0.80	0.87	0.82	0.33	0.31
TL 2	0.56	0.43	0.63	0.31	0.87	0.76	0.96	0.47	0.51
TL 3	0.54	0.44	0.69	0.42	0.67	0.80	0.89	0.67	0.72
TL 4	0.53	0.52	0.97	0.87	0.94	0.88	0.85	0.73	0.76
TL 5	0.51	0.61	0.79	0.77	0.91	0.92	0.87	0.67	0.72
TL 6	0.50	0.16	0.60	0.81	0.28	0.75	0.76	0.60	0.66
TL 7	0.48	0.28	0.55	0.34	0.40	0.84	0.80	1.00	0.87
TL 8	0.48	0.49	0.65	0.64	0.58	0.98	0.86	0.87	0.83
TL 9	0.50	0.70	0.63	0.56	0.77	0.80	0.95	0.67	0.72
TL 10	0.47	0.56	0.63	0.54	0.86	0.92	0.83	0.73	0.76
TL 11	0.48	0.52	0.58	0.25	0.74	0.91	0.96	0.67	0.72
TL 12	0.46	0.58	0.36	0.30	0.88	0.75	0.94	0.60	0.66
Min	0.46	0.16	0.36	0.25	0.28	0.75	0.76	0.33	0.31
Max	0.56	0.70	0.97	0.87	0.94	0.98	0.96	1.00	0.87
Average	0.50	0.49	0.65	0.56	0.72	0.85	0.87	0.67	0.68

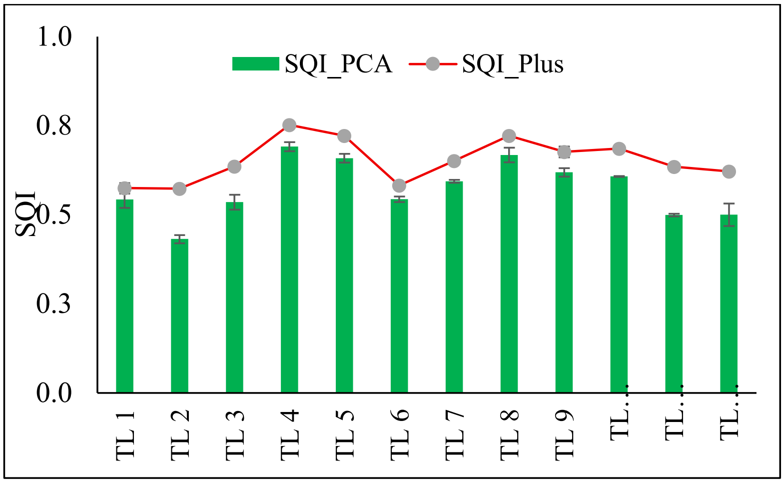
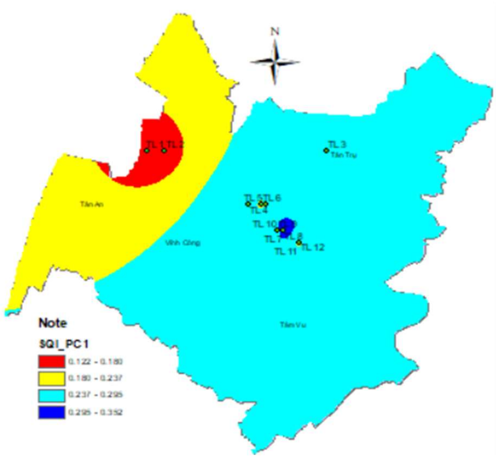
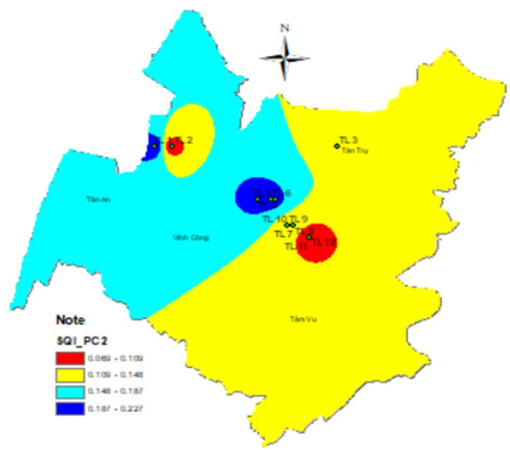


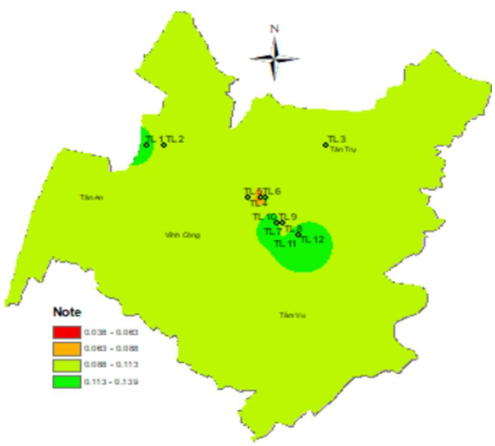
Fig. 3: SQI estimation results of soil samples by 2 methods (Additive method and PCA).



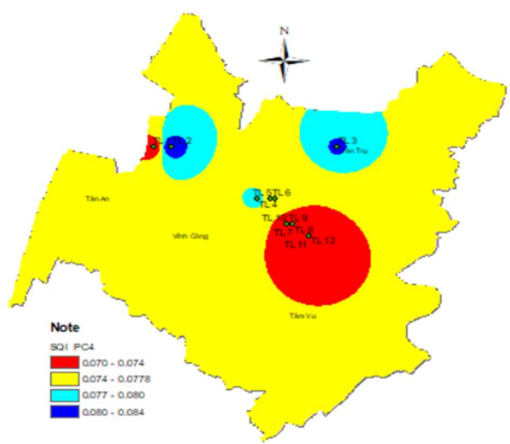
a. SQI for PC1



b. SQI for PC2



c. SQI for PC3



d. SQI for PC4

Fig. 4: SQI for each Principal Component.

due to the influence of highly correlated parameters causing multicollinearity in the calculation.

According to the classification scale used in Li et al.'s study, SQI_Plus classified three samples—TL 4, TL 5, and TL 8—as being in good condition (SQI > 0.7), representing 25%. Six samples fall into the average condition range (SQI 0.6 to 0.7), accounting for 50%, while three samples, TL 1, TL 2 and TL 6, are degraded (SQI 0.4 to 0.6), making up 25%. In contrast, SQI_PCA indicates no soil samples are in good condition, five are in the average range (41.6%), and seven are degraded (58.3%). These findings suggest that SQI_PCA is more appropriate for managing soil quality degradation risks. The soil quality across the studied farms tends to decline, possibly due to recent decreases in dragon fruit prices, which have reduced investments in soil care, improvement, and nutrition. Field observations also reveal that most farms have been cultivated for over a decade, and fragmented cultivation has resulted in a mix of soil qualities, both good and poor (Fig. 3). This variability poses a risk to sustainable agriculture. ble agriculture.

3.4. Mapping: Spatial Interpolation Analysis

Soil Quality Index (SQI) mapping was conducted using the IDW interpolation method in ArcGIS 10.8 to assess the spatial variation of soil quality across the study area based on SQI estimates. As shown in Fig. 4a, SQI_PC1 indicated that samples TL1 and TL2, located west of the dragon fruit plantations, had lower values than other farms, correlating with low clay content and signs of erosion caused by improper irrigation practices and overly strong hoses, as observed in the field. The SQI values for PC2 (Fig. 4b) revealed that the central zone (samples TL11, TL12) and the western zone (TL1) require management interventions, such as the addition of phosphorus and organic matter to enhance soil quality. In Fig. 4c, the SQI values for PC3 in the central (TL10, TL11, TL12) and western (TL1) regions were low, indicating the need to regulate the excessive use of mineral fertilizers like ammonium nitrogen. Instead, increasing organic matter to manage nitrogen levels is recommended. Finally, Fig. 4d demonstrated that the central and western parts of the study area exhibited low SQI values, signifying poor soil pH regulation that warrants closer attention.

Fig. 5 illustrates that the soil quality in the central growing region predominantly falls into the medium category, showing signs of degradation. In contrast, the western area exhibits poor soil quality, indicated by sites TL1 and TL2. These western samples correspond to recently converted dragon fruit cultivation zones where farmers are actively managing soil conditions. The central region also contains areas with lower soil quality, such as sample TL6 (see Fig. 5). The findings demonstrate that the Soil Quality Index (SQI) is influenced by soil properties, crop age, and

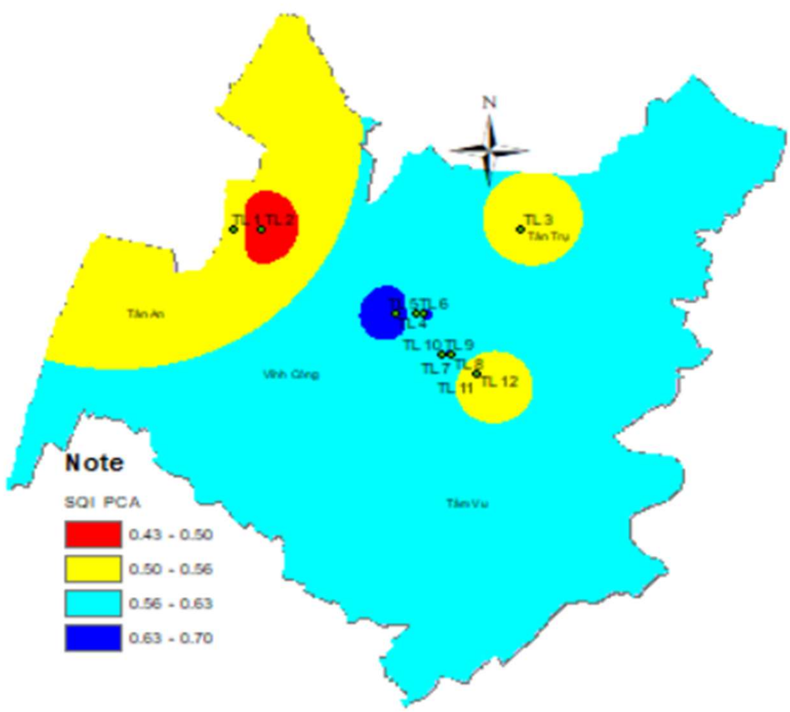


Fig. 5: Spatial distribution map of SQI values according to PCA in dragon fruit farms.

farmers' experience. Implementing a soil quality index distribution map can serve as a valuable tool for managers to identify soil degradation trends early.

4. CONCLUSIONS

Agricultural practices play a crucial role in shaping soil health by impacting its physical, chemical, and fertility attributes. This study assesses the soil quality index (SQI) based on the physico-chemical properties of soil in dragon fruit farms in Tây Ninh, Vietnam. Principal component analysis (PCA) combined with GIS mapping was used to visualize the spatial distribution of SQI. The minimum data set (MDS) derived from PCA explained 89.0% of the total variance, with four principal components contributing 33.2%, 22.7%, 19.9%, and 13.2%, respectively. Key parameters such as cation exchange capacity (CEC), available phosphorus (P_av), electrical conductivity (EC), and pH were selected as representative variables with weights of 0.37, 0.26, 0.22, and 0.15. SQI values calculated through PCA revealed improved risk management, with 41.6% of soils classified as high quality and 58.3% as degraded. These findings demonstrate that integrating PCA and GIS provides an effective approach for detecting soil degradation trends in areas of intensive perennial crop cultivation.

5. ACKNOWLEDGEMENTS

This research was conducted with the support of students majoring in Land Management, Course 16. We also express our deep gratitude to the reviewers for their in-depth comments, which contributed significantly to improving the manuscript.

6. REFERENCES

Abdel-Fattah, M.K., Mohamed, E.S., Wagdi, E.M., Shahin, S.A., Aldosari, A.A., Lasaponara, R. and Alnaimy, M.A., 2021. Quantitative evaluation of soil quality using Principal Component Analysis: The case study of El-Fayoum depression, Egypt. *Sustainability*, 13(4), p.1824. [DOI]

Abdu, A., Laekemariam, F., Gidago, G. and Getaneh, L., 2023. Explaining the soil quality using different assessment techniques. *Applied and Environmental Soil Science*, 2023. [DOI]

Aisah, B.N., Ikhwani, A.T., Setiawan, A.B., Purwantoro, A., Respatie, D.W., Ambarwati, E., Anggraeni, L., Teo, C.H., Setyawan, C. and Khamidah, A., 2025. Assessment of F_5 mungbean genotypes from intergeneric hybridization with common bean for agronomic performance, heritability, variance components, and genetic uniformity. *Caraka Tani: Journal of Sustainable Agriculture*, 40(3), pp.347-360. [DOI]

Baethgen, W. and 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), pp.961-969.

Bandyopadhyay, S. and 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, p.106351. [DOI]

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. and Brussaard, L., 2018. Soil quality – A critical review. *Soil Biology and Biochemistry*, 120, pp.105-125. [DOI]

Damiba, W.A.F., Gathenya, J.M., Raude, J.M. and 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), p.e27219.

George, J., 1962. Hydrometer method improved for making particle size analyses of soils. *Agron J*, 54, pp.464-465.

Hanif, S., Ali, S., Chaudhry, A.H., Sial, N., Marium, A. and Mehmood, T., 2025. A review on soil metal contamination and its environmental implications. *Nature Environment and Pollution Technology*, 25(1), Article ID D1684. [DOI]

Kahsay, A., Haile, M., Gebresamuel, G. and Mohammed, M., 2025. Developing soil quality indices to investigate degradation impacts of different land use types in Northern Ethiopia. *Heliyon*, 11(1), p.e41052.

Kawaguchi, K. and Kyuma, K., 1975. Paddy soils in tropical Asia: Part 3. Correlation and regression analyses of the soil data. *South East Studies*, 13(1), pp.97-118.

Li, Y., Zhao, H., Liu, J., Chaonan, C. and Yuxuan, G., 2024. A framework for selecting and assessing soil quality indicators for sustainable soil management in waste dumps. *Scientific Reports*, 14(1), p.8491. [DOI]

Mai, T.D., Dinh, H.H.D., Ha, N.S. and Nguyen, C.K., 2025. Expanding export markets for Vietnamese dragon fruit. *International Journal of Scientific Research and Management (IJSRM)*, 13(4), pp.245-260.

Mukherjee, A. and Lal, R.L., 2014. Comparison of soil quality index using three methods. *PLoS ONE*, 9(8), p.e105981. [DOI]

Plonsky, L., 2015. Advancing quantitative methods in second language research. In: *Routledge Handbook of Second Language Research*. Routledge, pp.145-165.

Rajput, R., Patley, S., Choudhary, B. and 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. *Journal of the Indian Society of Soil Science*, 71(3), pp.287-297.

Rangel-Peraza, J.G., Padilla-Gasca, E., López-Corrales, R., Medina, J.R., Bustos-Terrones, Y., Amabilis-Sosa, L.E., Rodríguez-Mata, A.E. and Osuna-Enciso, T., 2017. Robust soil quality index for tropical soils influenced by agricultural activities. *Journal of Agricultural Chemistry and Environment*, 6, pp.199-221.

Țopa, D.-C., Căpșună, S., Calistru, A.-E. and Ailincăi, C., 2025. Sustainable practices for enhancing soil health and crop quality in modern agriculture: A review. *Agriculture; Basel*, 15(9), p.945.

Van, P.N., 2025. Assessing soil health through multivariate analysis: A focus on durian cultivation in Cho Lach, Ben Tre Province, Vietnam. *Vietnam Journal of Agricultural Sciences*, 8(2), pp.115-128.