Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Assessing and Mapping Soil Quality in Northern Vietnam: A Minimum Data Set Approach Comparing the Northern Mountainous and Midland Region and the Red River Delta

Lim Duong Thi
Vietnam Academy of Science and Technology
Quan Dang Tran
Vietnam Academy of Science and Technology
Hue Nguyen Thi
Vietnam Academy of Science and Technology
Huong Nguyen Thi Lan
Vietnam Academy of Science and Technology
Dung Le Thi Phuong
Vietnam Academy of Science and Technology
Thuy Nguyen Thi Huong
Vietnam Academy of Science and Technology
Thang Cu Sy
Vietnam Academy of Science and Technology
Thu Nguyen Thi
Vietnam Academy of Science and Technology
Cuong Nguyen Viet
Vietnam Academy of Science and Technology
Trang Nguyen Thi Thu
Ministry of Agriculture and Environment
Hung Nguyen Duc
Ministry of Agriculture and Environment
Quy Nguyen Ngoc
Ministry of Science and Technology
Toan Nguyen-Sy
nstoan@ute.udn.vn

University of Da Nang

Research Article

Keywords: SQI, MDS, PCA, entropy weighting, spatial interpolation, northern Vietnam

Posted Date: September 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10906736/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Submision to Journal of Environmental Geochemistry and Health as original research entittled,

Assessing and Mapping Soil Quality in Northern Vietnam: A Minimum Data Set Approach Comparing the Northern Mountainous and Midland Region and the Red River Delta

Duong Thi Lim • Dang Tran Quan • Nguyen Thi Hue • Nguyen Thi Lan Huong • Le Thi Phuong Dung • Nguyen Thi Huong Thuy • Cu Sy Thang • Nguyen Thi Thu • Nguyen Viet Cuong • Nguyen Thi Thu Trang • Nguyen Duc Hung • Nguyen Ngoc Quy • Nguyen-Sy Toan*

D.T. Lim • D.T. Quan • N.T. Hue • N. T. L. Huon • L. T. P. Dung • N. T. H. Thuy • C. S. Thang • N. T. Thu • N. V. Cuong
Institute of Earth Sciences - Viet Nam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Ha Noi, Viet Nam

N. T. T. Trang • N. D. Hung
National Center for Land Survey and Planning - Ministry of Agriculture and Environment

N. N. Quy
National Academy for Advanced Technology and Innovation - Ministry of Science and Technology

N-S. Toan (✉)
Faculty of Chemical Technology and Environment, The University of Da Nang-University of Technology and Education, 48 - Cao Thang, Hai Chau, Danang, Viet Nam
Email: nstoan@ute.udn.vn)

Abstract

Soil quality assessment is essential for understanding spatial variability in soil conditions and supporting sustainable soil management. This study aimed to assess and spatially characterize soil quality in the Northern Mountainous and Midland Region (NMMR) and the Red River Delta (RRD) of northern Vietnam using an integrated Principal Component Analysis (PCA) Minimum Data Set (MDS) entropy-weighting approach. A total of 346 soil samples were collected, including 261 from the NMMR and 85 from the RRD, and analyzed for physicochemical properties and potentially toxic elements. PCA was used to identify the major dimensions of soil variability and reduce redundancy among indicators, resulting in an MDS of seven indicators: Pb, As, bulk density, Cd, Zn, organic carbon, and clay. The selected indicators were weighted using the Shannon entropy method and integrated into a Soil Quality Index (SQI). The mean SQI was significantly higher in the RRD (0.4633 ± 0.0537) than in the NMMR (0.4280 ± 0.0629), representing an 8.25% difference with a moderate effect size (Cohen's $d = 0.581$, $p < 0.001$). The NMMR also exhibited a wider range of SQI values, indicating greater within-region variability. For spatial prediction, Inverse Distance Weighting (IDW) with a power parameter of 1.5 performed best in the NMMR, whereas IDW with a power parameter of 1.0 performed best in the RRD based on cross-validation RMSE. However, the relatively low cross-validation R^2 values indicated limited spatial predictability from geographic proximity alone. As conclusion, the integrated PCA-MDS entropy framework provided a practical regional-scale approach for assessing soil quality and identifying areas requiring further investigation and management.

Keywords: SQI; MDS; PCA; entropy weighting; spatial interpolation; northern Vietnam.

Introduction

Soil is a finite, slowly renewable resource that underpins food security, water regulation, and ecosystem services, yet it is increasingly threatened by degradation processes such as compaction, organic matter depletion, and contamination by potentially toxic elements (PTEs) (Almasoud et al., 2014; Anh et al., 2026; Toan et al., 2025; Tuan Anh et al., 2026). Because soil quality cannot be measured directly, it is commonly assessed using composite indices that integrate multiple physical, chemical, and biological indicators into a single, interpretable value. Among these, the Soil Quality Index (SQI) derived from a Minimum Data Set (MDS) has become one of the most widely applied frameworks, as it reduces a large pool of candidate indicators to a smaller, non-redundant subset selected through multivariate statistical techniques such as principal component analysis (PCA), each indicator subsequently standardized to a unitless score and combined using a weighting scheme (Maaz et al., 2023; Raiesi, 2017; Stenberg, 1999; Yang et al., 2026; P. Yu et al., 2018). A critical, and often inconsistently applied, step in the MDS approach is the derivation of indicator weights. While earlier studies frequently assigned weights based on the proportion of variance explained by each retained principal component, this approach can be sensitive to the specific PCA solution and does not always reflect the true information content of each indicator across the sampled population (Wang et al., 2025; Yang et al., 2026). The Shannon entropy method offers an alternative, data-driven weighting scheme that assigns greater weight to indicators with higher variability (greater discriminatory power) among sampling sites, and has been increasingly adopted in recent soil quality assessments as a more objective alternative to expert opinion or variance-based weighting (Borchers & Pieler, 2010; Ghosh et al., 2025; Xie et al., 2018).

Beyond inherent physical and chemical fertility, PTEs such as Cd, Pb, As, Cr, and Zn represent a growing concern for soil quality and, by extension, human and ecosystem health, given their

persistence, potential for bioaccumulation, and the multiple anthropogenic and geogenic pathways through which they enter agricultural soils, including mining activity, industrial emissions, traffic, agrochemical use, and parent material weathering (Anh et al., 2026; Cai et al., 2019; Kutakimina et al., 2026; T. Nguyen-Sy et al., 2025). Incorporating PTEs alongside conventional physical and chemical indicators within a single SQI framework therefore provides a more health-relevant and environmentally comprehensive assessment of soil condition than fertility indicators alone (Jahany & Rezapour, 2020; Mamehpour et al., 2021; Rahmanipour et al., 2014; Sultanovna & Ismailovich, 2026; Yeilagi et al., 2021).

Vietnam presents a particularly informative setting for comparative soil quality assessment, as its northern territory encompasses two agroecologically contrasting regions: the Northern Mountainous and Midland Region (NMMR), characterized by steep terrain, heterogeneous geology, and mixed upland cultivation (Anh et al., 2026; Clemens et al., 2010; Tuan Anh et al., 2026), and the Red River Delta (RRD), a low-lying alluvial plain under long-term intensive paddy rice cultivation (Mussgnug et al., 2006; Phuong et al., 2009). These two regions differ markedly in parent material, land-use history, agrochemical input, and erosion/deposition dynamics, all of which are expected to shape both the fertility status and the contamination status of their soils as our recent research (Anh et al., 2026; Tuan Anh et al., 2026). Despite this contrast, direct, quantitative comparisons of soil quality between the two regions using a unified indicator framework remain limited, and few studies have combined an entropy-weighted MDS-based SQI with geospatial interpolation to characterize the spatial distribution of soil quality across such a large and heterogeneous area.

Due to the gap of above research, this study therefore aimed to (i) characterize and statistically compare the physical, chemical, and potentially toxic element properties of soils in the NMMR and RRD regions of northern Vietnam; (ii) identify a Minimum Data Set of soil quality indicators using principal component analysis and derive their weights using the Shannon

entropy method; (iii) compute a composite SQI for each sampling site and compare SQI values between the two regions; and (iv) evaluate spatial interpolation methods and map the spatial distribution of soil quality across both regions to identify priority areas for soil management.

Materials and Methods

Study area and soil sampling

Soil sampling was conducted in two contrasting agroecological regions of northern Vietnam: the Northern Mountainous and Midland Region (NMMR) and the Red River Delta (RRD) (Fig. 1).

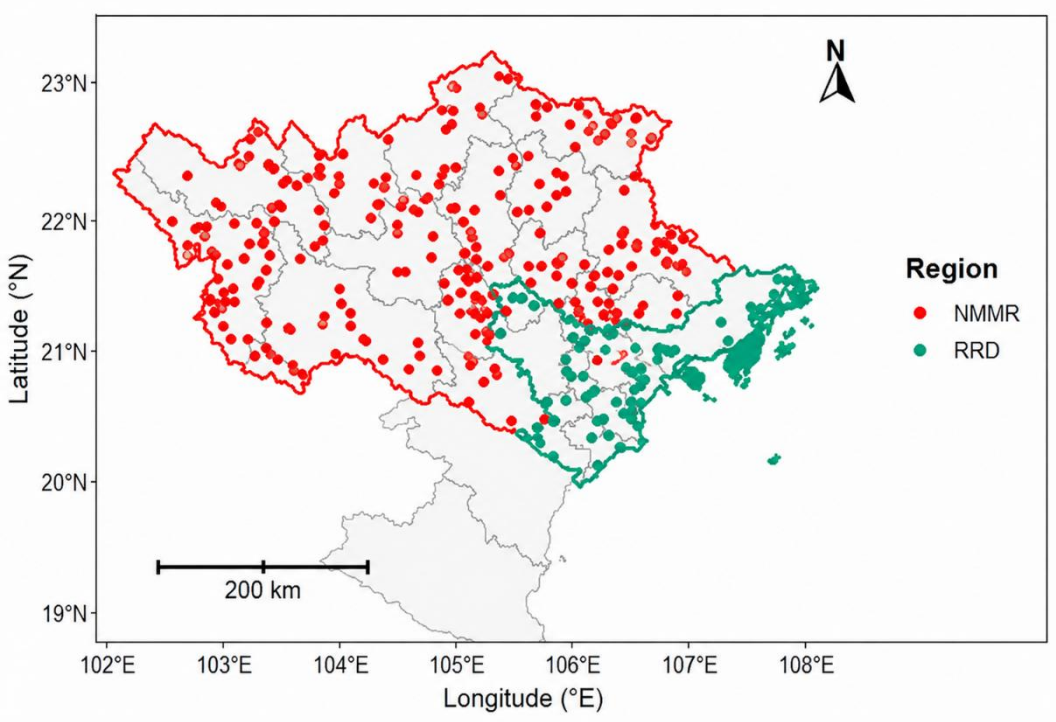


Fig. 1 Study area and spatial distribution of soil sampling sites in the NMMR and RRD

The study area covered approximately 19 - 23°N latitude and 102 - 108°E longitude. A total of 346 soil samples were collected, including 261 samples from the NMMR and 85 samples from the RRD. The sampling sites were georeferenced using a handheld Global Positioning System (GPS) receiver. Field investigation and sampling were conducted in accordance with Circular No. 11/2024/TT-BTNMT, issued by the Ministry of Natural Resources and

Environment of Vietnam on 31 July 2024, which sets out technical requirements for soil investigation and assessment, as well as for soil protection, improvement, and restoration. Sampling locations were established based on relatively homogeneous soil and terrain units. The delineation of sampling units considered at least two relatively homogeneous factors, namely soil type and terrain. Depending on the degree of homogeneity and complexity of soil and terrain distribution, sampling density was approximately one sampling location per 2,500 ha for areas with highly homogeneous soil and terrain, one per 2,000 ha for relatively homogeneous areas, and one per 1,500 ha for areas with complex and interspersed soil and terrain conditions. Soil samples were collected from the surface soil layer at a depth of 0–50 cm, depending on the thickness of the surface soil horizon. Approximately 2 kg of soil was collected at each sampling location. Sampling and laboratory analyses were conducted from January to April 2025.

Soil physicochemical and potentially toxic element analyses

A total of 18 soil quality indicators were determined for the collected samples, covering physical properties, chemical properties, nutrient status, and potentially toxic elements (PTEs). The indicators included bulk density, pH measured in KCl (pH_KCl), sand, silt, clay, cation exchange capacity (CEC), organic carbon (C), total nitrogen (N), total P_2O_5, total K_2O, and eight PTEs (Cd, Cu, As, Pb, Cr, Zn, Ni, and Co). Soil particle-size distribution was determined using the pipette method. Bulk density was determined from the mass of oven-dried soil per unit volume of the undisturbed soil core. Particle density and porosity were determined from soil mass and volume measurements. Soil pH_KCl was measured after extraction with 1 N KCl at a soil-to-solution ratio of 1:5 using a calibrated pH meter. Organic carbon was determined using the Walkley - Black method, while total nitrogen was determined using the Kjeldahl method. Total P_2O_5 was determined by a colorimetric method, total K_2O by flame photometry, and CEC by ammonium acetate extraction at pH 7. Concentrations of Cd, Cu, As, Pb, Cr, Zn, Ni, and Co were determined as potentially toxic element indicators. The analytical quality-control procedure should be described using the actual digestion and instrumental determination methods used for the soil samples. The available Certificate of Analysis documents a certified freshwater-sediment reference material (CRM016-50G), with certified values for As, Cd, Cr, Co, Cu, Pb, Ni, and Zn, among other elements. The certificate indicates that the CRM is suitable for digestion using USEPA 3050, 3051, or equivalent methods.

Analytical framework for soil quality assessment

The overall analytical framework adopted in this study is presented in Fig. 2.

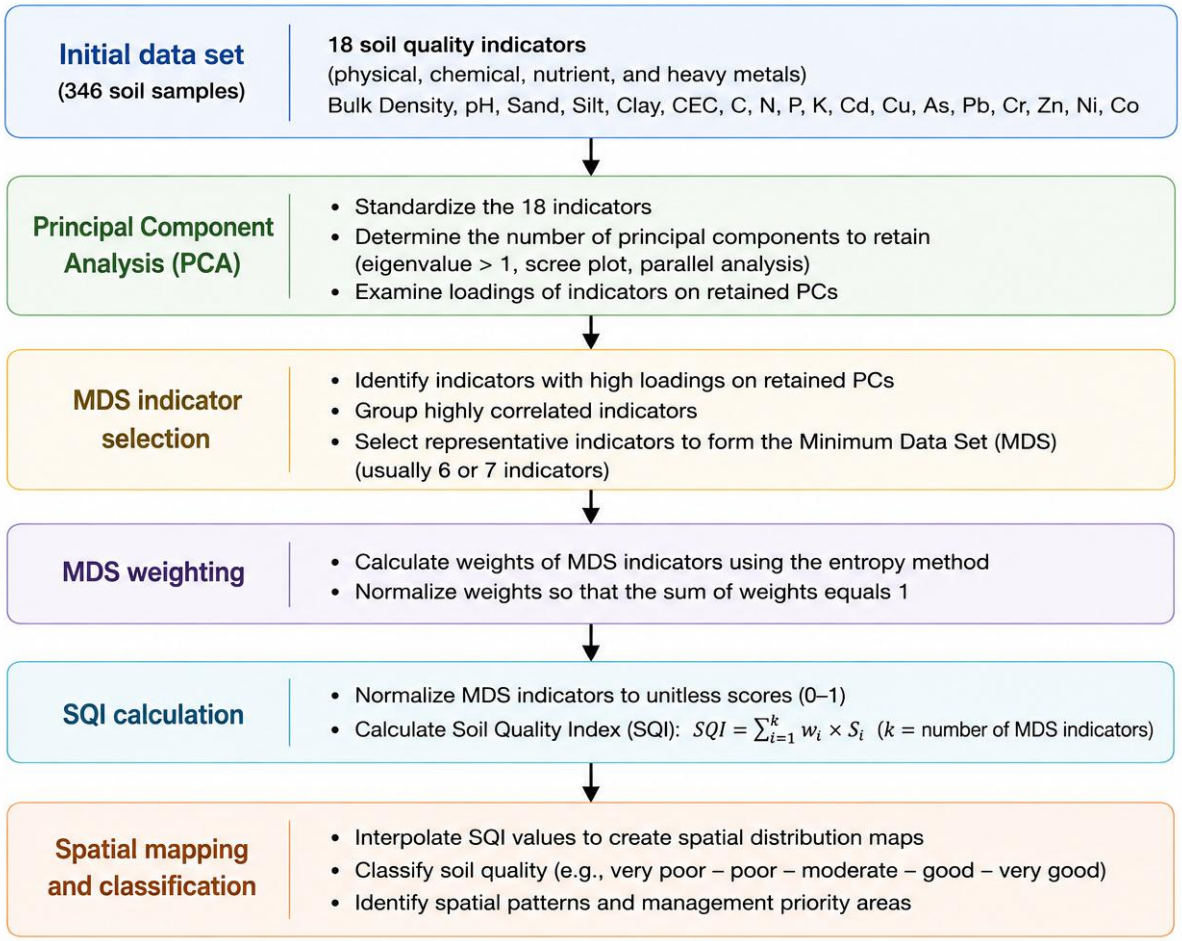


Fig. 2 Workflow of the soil quality assessment framework

The assessment consisted of a sequential workflow involving soil sampling and laboratory analysis, data preprocessing, principal component analysis (PCA), Minimum Data Set (MDS) selection, entropy-based weighting, SQI calculation, spatial interpolation, and soil-quality mapping. Initially, 18 soil quality indicators were measured for the collected soil samples. Following data preprocessing, the candidate indicators were evaluated using PCA to identify the major dimensions of soil-quality variation and to reduce redundancy among indicators. A subset of representative indicators was subsequently selected to form the MDS. The selected MDS indicators were assigned data-driven weights using the Shannon entropy method and transformed into dimensionless scores according to their respective scoring functions. The weighted scores were then integrated to calculate the SQI for each sampling location.

Subsequently, the spatial distribution of SQI was evaluated separately for the NMMR and RRD using alternative spatial interpolation approaches. Cross-validation was used to assess interpolation performance, and the best-performing model for each region was used to generate the final spatial SQI surface. The resulting maps were then classified into soil-quality categories to facilitate regional interpretation and identification of areas requiring further investigation or management attention.

Principal component analysis and Minimum Data Set selection

Principal component analysis (PCA) was used to reduce the dimensionality of the soil-quality dataset and identify a Minimum Data Set (MDS) for subsequent SQI calculation. Of the 18 initially measured soil quality indicators, silt content was excluded before PCA, resulting in 17 candidate indicators for multivariate analysis. Geographic coordinates were not considered soil quality indicators and were therefore excluded from the PCA. PTE concentrations were $\log_{10}$-transformed before PCA to reduce the influence of highly right-skewed distributions. The number of principal components (PCs) retained for subsequent indicator selection was determined using three complementary criteria: the eigenvalue-greater-than-one criterion, the scree plot, and parallel analysis.

Entropy-based weighting of MDS indicators

The relative weights of the seven MDS indicators were determined using the Shannon entropy weighting method. Entropy weighting was used to quantify the relative contribution of each indicator to the overall information based on its variability across sampling sites. The resulting weights were normalized so that the sum of all indicator weights was equal to 1.

Soil Quality Index calculation

The selected MDS indicators were transformed into dimensionless scores ranging from 0 to 1 using nonlinear scoring functions. The scoring direction was determined according to the expected relationship between each indicator and soil quality. Pb, As, Cd, and Zn were evaluated using a “lower is better” function, whereas C was evaluated using a “higher is better” function. Bulk density and clay were evaluated using optimum-response functions

according to their respective relationships with soil quality. The SQI for each sampling location was calculated using a weighted additive approach:

$$SQI = \sum_{i=1}^{n} w_i S_i$$

where SQI is the Soil Quality Index, w_i is the normalized entropy-derived weight of the ith MDS indicator, S_i is the corresponding dimensionless score ranging from 0 to 1, and n is the number of indicators included in the MDS ($n = 7$). Thus, the SQI integrates the standardized performance and relative information contribution of the seven selected indicators into a single dimensionless measure.

Statistical and spatial analysis

Descriptive statistics, including mean, standard deviation, median, minimum, and maximum, were calculated for each soil quality indicator separately for the NMMR and RRD. Differences in soil physicochemical properties and potentially toxic element concentrations between the two regions were evaluated using Welch's *t*-test, which does not assume equal variances between groups. Statistical significance was assessed at $p < 0.05$. For the SQI, regional differences were assessed using the mean difference, percentage difference, Cohen's *d* effect size, Welch's *t*-statistic, corresponding degrees of freedom, *p*-value, and 95% confidence interval (CI) for the difference between regional means. All statistical analyses, PCA, entropy weighting, SQI calculations, spatial interpolation, cross-validation, and graphical outputs were performed using R version 3.5.4 (R Core Team, Vienna, Austria).

Results and Discussion

Soil physicochemical properties and potentially toxic elements

The descriptive statistics of soil physicochemical properties and potentially toxic elements (PTEs) in the Northern Mountainous and Midland Region (NMMR) and Red River Delta (RRD) are presented in Table 1. Substantial variability was observed among the investigated

soil indicators, particularly for several PTEs. Significant differences between the two regions were observed for bulk density, total nitrogen (N), and chromium (Cr), whereas the other measured indicators did not differ significantly at $p < 0.05$.

Table 1. Descriptive statistics of soil physicochemical properties and potentially toxic elements in the NMMR and RRD regions

Group	Variable	Unit	NMMR (Mean ± SD)	RRD (Mean ± SD)	p-value
Physical properties	Bulk density	g cm^{-3}	1.409 ± 0.098	1.322 ± 0.104	<0.001
Physical properties	pH (KCl)	-	4.809 ± 1.094	5.004 ± 1.262	0.205
Physical properties	Sand	%	54.563 ± 14.116	53.118 ± 19.958	0.537
Physical properties	Clay	%	19.162 ± 9.908	17.981 ± 10.025	0.346
Physical properties	CEC	meq 100 g^{-1}	10.623 ± 4.482	10.418 ± 4.964	0.735
Chemical properties	C	%	1.575 ± 0.916	1.698 ± 0.904	0.281
Chemical properties	N	%	0.120 ± 0.066	0.139 ± 0.076	0.036
Chemical properties	P_2O_5	%	0.180 ± 0.658	0.189 ± 0.160	0.845
Chemical properties	K_2O	%	0.879 ± 0.614	0.984 ± 0.672	0.205
Potentially toxic elements	Cd	mg kg^{-1}	0.469 ± 2.220	0.198 ± 0.182	0.052
Potentially toxic elements	Cu	mg kg^{-1}	64.817 ± 552.856	29.914 ± 32.396	0.311
Potentially toxic elements	As	mg kg^{-1}	35.793 ± 342.067	10.600 ± 18.673	0.237
Potentially toxic elements	Pb	mg kg^{-1}	56.861 ± 349.385	23.881 ± 15.933	0.130
Potentially toxic elements	Cr	mg kg^{-1}	64.754 ± 207.530	28.625 ± 34.795	0.007
Potentially toxic elements	Zn	mg kg^{-1}	116.445 ± 797.352	50.446 ± 38.887	0.184
Potentially toxic elements	Ni	mg kg^{-1}	32.781 ± 169.595	14.519 ± 12.707	0.086
Potentially toxic elements	Co	mg kg^{-1}	16.903 ± 75.031	9.977 ± 14.579	0.159

Bulk density was significantly higher in the NMMR than in the RRD, with mean values of 1.409 ± 0.098 and 1.322 ± 0.104 g cm^{-3}, respectively ($p < 0.001$). This difference indicates a generally denser soil condition in the NMMR. In contrast, the RRD exhibited a slightly higher mean organic carbon concentration (1.698 ± 0.904%) than the NMMR (1.575 ± 0.916%), although the difference was not statistically significant ($p = 0.281$). The mean clay contents were also similar between the NMMR and RRD (19.162 ± 9.908% and 17.981 ± 10.025%, respectively; $p = 0.346$), as were pH_KCl and CEC.

Total N was significantly higher in the RRD than in the NMMR (0.139 ± 0.076% versus 0.120 ± 0.066%; $p = 0.036$). The higher N concentration in the RRD may reflect differences in agricultural management and nutrient inputs; however, the available dataset does not permit direct testing of these potential drivers (Mussgnug et al., 2006). It is noted that P_2O_5 and K_2O did not differ significantly between regions (p-values of 0.845 and 0.205, respectively), whereas PTE concentrations showed pronounced variability, particularly in the

NMMR. For example, the standard deviations of Cu, As, Pb, Cr, and Zn in the NMMR were 552.856, 342.067, 349.385, 207.530, and 797.352 mg kg^{-1}, respectively, considerably exceeding their corresponding means. Similar but less pronounced variability was observed in the RRD (Phuong et al., 2008; Tuan Anh et al., 2026). Such large dispersion indicates that the PTE distributions were highly heterogeneous and likely strongly skewed, supporting the use of logarithmic transformation before multivariate analysis. Among the PTEs, Cr showed a statistically significant regional difference, with a substantially higher mean concentration in the NMMR than in the RRD (64.754 ± 207.530 versus 28.625 ± 34.795 mg kg^{-1}; p = 0.007). In contrast, Cd, Cu, As, Pb, Zn, Ni, and Co did not differ significantly between the two regions ($p > 0.05$). We found that Cd approached the conventional significance threshold ($p = 0.052$), whereas Ni showed a weaker tendency toward regional differentiation (p = 0.086). Although the NMMR generally exhibited higher mean concentrations of the investigated PTEs than the RRD, these differences were not statistically significant for most elements. For instance, the mean Pb and As concentrations in the NMMR were approximately 2.4- and 3.4-fold those in the RRD, respectively, but the corresponding p-values were 0.130 and 0.237. This highlights an important distinction between differences in regional means and statistically supported regional differences (Anh et al., 2026; T. Nguyen-Sy et al., 2025; Tuan Anh et al., 2026). The large within-region variability, particularly in the NMMR, likely contributed to the lack of statistical significance for several PTEs. In general, the univariate results indicate that the two regions shared broadly comparable soil chemical characteristics. At the same time, the most evident regional differences were associated with bulk density, total N, and Cr. The large variability in PTE concentrations further suggests that soil quality may be controlled by substantial local-scale heterogeneity rather than by regional differences alone.

PCA and MDS selection

Principal component analysis was used to reduce the dimensionality of the soil-quality dataset and to identify indicators that contribute most strongly to the major patterns of variation. Four principal components were retained based on the PCA screening procedure, including eigenvalues, the scree plot, and parallel analysis (Fig. S1). The first two components accounted for 39.4% and 20.7% of the total variance, respectively, corresponding to 60.1% of the variance explained by PC1 and PC2 together. The PCA loading matrix in Table 2 shows that the selected indicators were distributed across different

components. Cd exhibited the highest loading among the variables included in the MDS, with a loading of 0.500 on PC4. Clay had a maximum loading of 0.441 on PC3, while As and Pb had maximum loadings of 0.422 and 0.409, respectively, on PC4. Organic C and bulk density showed relatively high loadings on PC2, with values of 0.404 and -0.402, respectively. Zn had a maximum loading of 0.359 on PC1.

Table 2. PCA loadings and selection of soil quality indicators for the Minimum Data Set*.

Variable	PC1	PC2	PC3	PC4	Max Loading	MDS selected
Cd	0,23	-0,133	-0,065	0,5	0,5	Yes
Clay	0,076	0,341	0,441	-0,12	0,441	Yes
N	0,135	0,431	-0,342	0,127	0,431	No
As	0,271	-0,086	0,163	0,422	0,422	Yes
Pb	0,281	-0,137	0,136	0,409	0,409	Yes
C	0,135	0,404	-0,349	0,075	0,404	Yes
Bulk density	-0,144	-0,402	0,03	0,004	0,402	Yes
Sand	-0,151	-0,373	-0,396	0,072	0,396	No
pH (KCl)	0,144	-0,175	-0,36	0,036	0,36	No
Zn	0,359	-0,141	0,027	-0,032	0,359	Yes
Ni	0,358	-0,133	-0,017	-0,255	0,358	No
Cu	0,345	-0,115	0,079	-0,203	0,345	No
Co	0,336	-0,148	-0,008	-0,309	0,336	No
Cr	0,322	-0,097	0,04	-0,245	0,322	No
P_2O_5	0,138	0,12	-0,296	0,07	0,296	No
K_2O	0,157	0,186	0,277	0,231	0,277	No
CEC	0,23	0,165	-0,238	-0,204	0,238	No

*Note: PCA excluded geographic coordinates and Silt. PTEs were log10-transformed before PCA. The MDS indicators and weights are based on the PCA/MDS results previously obtained.

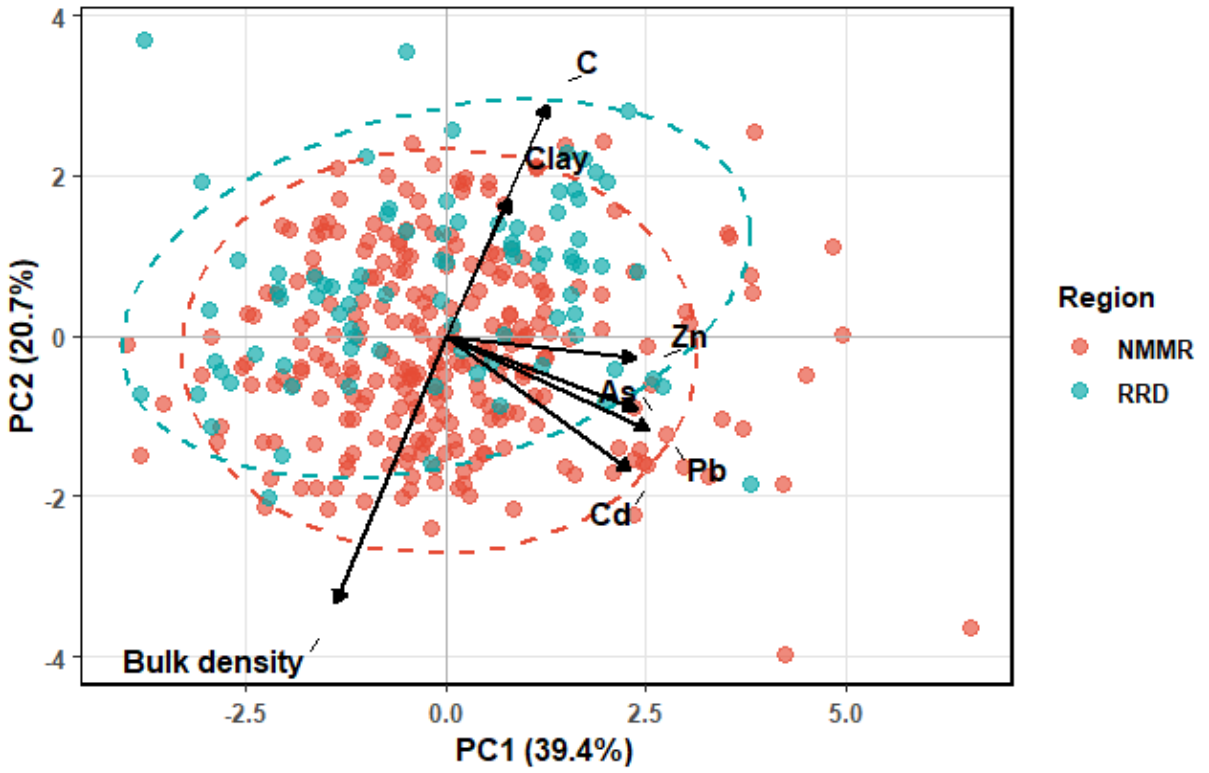


Fig. 3 Principal component analysis biplot showing the relationships between soil quality indicators and sampling sites in the NMMR and RRD regions.

The PCA biplot provides a visual representation of these relationships (Fig. 3). The vectors representing Cd, Pb, and As point broadly in similar directions, indicating a positive association among these PTE-related variables in the multivariate structure. Zn also lies in the same general direction, although its vector is shorter. In contrast, bulk density is oriented in an opposing direction along PC2, whereas C and clay are primarily associated with the positive PC2 direction. These patterns suggest that the principal sources of variation in the soil dataset were not represented by a single soil property but by several groups of correlated characteristics (Luo et al., 2017; Toan Nguyen-Sy et al., 2025; Pei et al., 2025). The association among Cd, As, Pb, and Zn indicates that PTE-related variability is an important dimension of soil quality variation. Meanwhile, the contrasting orientation of bulk density and the positive association of C and clay with PC2 indicate that physical and organic-matter characteristics provide an additional dimension of soil condition. Based on the loading criteria and redundancy considerations, seven indicators (Pb, As, bulk density, Cd, Zn, C, and clay) were selected to form the MDS. Notably, some variables that exhibited significant regional differences were not selected. For example, N and Cr differed significantly between the NMMR and RRD, but neither was included in the MDS. Conversely, Pb, As, Cd, and Zn

were retained despite the absence of statistically significant regional differences for most of these elements.

This distinction is important because MDS selection is intended to capture the principal structure of variation and reduce redundancy, rather than select variables showing significant differences between regions. This principle is consistent with previous SQI studies, in which PCA-based MDS selection was used to retain representative indicators while minimizing redundancy among correlated soil properties (Raiesi, 2017; Yang et al., 2026; Pujia Yu et al., 2018). Therefore, the MDS should not be interpreted as a list of the indicators that differ most strongly between NMMR and RRD. Instead, it represents a reduced set of indicators intended to preserve the major dimensions of soil-quality information. The substantial overlap between the NMMR and RRD observations in the PCA biplot further indicates that the two regions do not form completely separated soil-quality groups. This supports the later finding that regional differences in SQI coexist with considerable within-region variability.

Entropy-derived weighting of the Minimum Data Set indicators

The entropy-derived weights of the seven MDS indicators are presented in Table 3. Pb received the highest weight (17.03%), followed by As (15.76%), bulk density (15.42%), Cd (15.18%), Zn (14.97%), organic C (13.58%), and clay (8.06%).

Table 3. Entropy-derived weights and scoring directions of the selected Minimum Data Set indicators*

MDS indicator	Raw weight	Weight (%)	Indicator type	Scoring direction*	Role in SQI
Pb	1,7266	17,03	Potentially toxic element	Lower is better	PTE-related soil quality
As	1,5981	15,76	Potentially toxic element	Lower is better	PTE-related soil quality
Bulk density	1,5632	15,42	Physical property	Optimum / lower is better	Soil physical condition
Cd	1,5389	15,18	Potentially toxic element	Lower is better	PTE-related soil quality
Zn	1,5174	14,97	Potentially toxic element	Lower is better	PTE-related soil quality
C	1,3765	13,58	Chemical property	Higher is better	Soil organic C status
Clay	0,8175	8,06	Physical property	Optimum / context dependent	Soil texture / retention

** Final scoring function/direction follow the exact scoring curves used in the SQI calculation. The table labels for Bulk density. Weight (%) = Raw weight / Σ(Raw weights) × 100. The seven MDS indicators account for 100% of the SQI weighting scheme. Clay presented as general interpretation rather than a replacement for the actual scoring formula.*

The four PTE indicators (Pb, As, Cd, and Zn) collectively accounted for approximately 63% of the total SQI weighting. The remaining approximately 37% was contributed by bulk

density, organic C, and clay. Thus, PTE-related indicators constituted the largest component of the weighting structure, while physical and organic characteristics provided complementary information on soil condition. Pb received the highest individual weight, indicating that it contributed the most discriminatory information under the entropy-based weighting procedure. Clay had the lowest weight (8.06%), suggesting that its variability contributed less information than that of the other selected indicators under the entropy framework. Entropy weighting was used to quantify the relative contribution of each indicator to information based on its variability across sampling sites (Ghosh et al., 2025; Xie et al., 2018). An entropy-derived weight reflects the information contribution and variability of an indicator within the specific dataset; it does not represent an absolute ranking of environmental toxicity or ecological importance (Abdul-Wahab et al., 2026; Guo et al., 2026; Uymaz et al., 2026). Therefore, the relatively high combined weight of PTEs indicates their importance within the statistical structure of this dataset rather than demonstrating that PTEs alone determine overall soil quality. The scoring directions also reflect the intended environmental interpretation of the selected indicators. Pb, As, Cd, and Zn were assigned a lower-is-better direction, organic C a higher-is-better direction, and bulk density and clay optimum/context-dependent responses. The combination of these scoring directions allows the SQI to integrate both potential contamination pressure and beneficial soil-quality characteristics.

Soil Quality Index and regional comparison

The calculated SQI differed significantly between the NMMR and RRD (Table 4; Fig. 4). The mean SQI was 0.4280 ± 0.0629 in the NMMR and 0.4633 ± 0.0537 in the RRD. The RRD therefore had a mean SQI that was 0.0353 units, or 8.25%, higher than the NMMR's.

Table 4. Comparison of soil quality index values between the NMMR and RRD regions

Statistic	NMMR	RRD
n	261	85
Mean ± SD	0.4280 ± 0.0629	0.4633 ± 0.0537
Median	0.4280	0.4691
Minimum	0.0739	0.3221
Maximum	0.7678	0.6394
Mean difference (RRD - NMMR)		0.0353
Percentage difference		8.25%
Cohen's d		0.581
Welch's t		-5.040
df		165.03
p-value		<0.001
95% CI (RRD − NMMR)		0.0215 to 0.0491

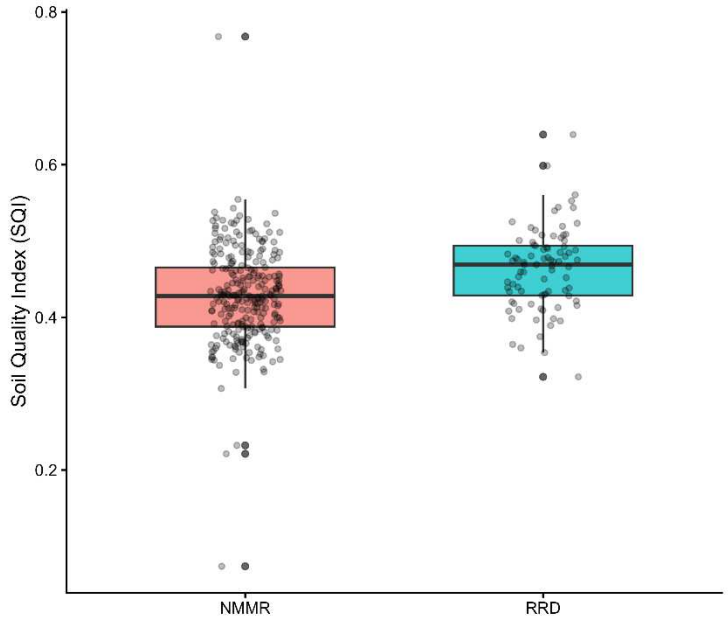


Fig. 4 Comparison of soil quality index values between the NMMR and RRD regions.

Welch's *t*-test confirmed that the difference between the two regional means was statistically significant (t = -5.040, df = 165.03, $p < 0.001$). The 95% confidence interval for the difference (RRD − NMMR) ranged from 0.0215 to 0.0491, indicating that the observed higher SQI in the RRD was unlikely to result from sampling variation alone. The Cohen's *d* value of 0.581 indicates a moderate standardized difference between the two regional distributions. The boxplot in Fig. 4 illustrates regional differences in SQI and reveals substantial overlap between the two distributions. The median SQI values were 0.4280 for the

NMMR and 0.4691 for the RRD. The NMMR exhibited a considerably wider range, from 0.0739 to 0.7678, compared with 0.3221 - 0.6394 in the RRD.

The wider SQI range in the NMMR indicates greater within-region heterogeneity. In particular, several NMMR observations had SQI values substantially below the regional mean, while others were considerably higher. In contrast, the RRD distribution was more concentrated around its central tendency. This pattern suggests that although the RRD had a higher average soil-quality index, regional averages alone do not capture the full spatial and local variability of soil quality. The higher SQI in the RRD may be partly consistent with its significantly lower bulk density, as bulk density was included in the MDS. However, the SQI is a composite index comprising seven indicators; therefore, the regional difference cannot be attributed to bulk density alone. Moreover, the RRD did not show significant differences for most of the other selected indicators. The observed SQI difference should therefore be interpreted as the integrated outcome of the selected indicators and their entropy-derived weights (Pujia Yu et al., 2018). An important finding is that the SQI result does not simply reproduce the results of the individual-indicator statistical comparisons. For example, Cr and total N showed significant regional differences in Table 1, but neither was included in the MDS. Conversely, Pb, As, Cd, Zn, C, and clay contributed to the SQI despite most of them showing no statistically significant regional difference. This demonstrates the value of an MDS-based SQI for integrating multiple dimensions of soil condition rather than relying solely on individual indicators or univariate significance tests, consistent with previous soil-quality assessments (Mamehpour et al., 2021; Rahmanipour et al., 2014; Yeilagi et al., 2021).

Spatial interpolation performance

The performance of the tested spatial interpolation models is presented in Table 5. Five IDW power parameters (p = 1.0, 1.5, 2.0, 2.5, and 3.0) were evaluated separately for the NMMR and RRD, alongside Ordinary Kriging. Model selection was based primarily on the lowest RMSE obtained through cross-validation.

Table 5. Cross-validation performance of spatial interpolation methods for the Soil Quality Index.

Region	Method	Parameter	RMSE	MAE	ME	R^2	Selected
NMMR	IDW	p = 1.0	0.05761	0.04148	0.00057	0.1568	
NMMR	IDW	p = 1.5	0.05732	0.04153	0.00089	0.1655	Yes
NMMR	IDW	p = 2.0	0.05953	0.04364	0.00088	0.0997	
NMMR	IDW	p = 2.5	0.06238	0.04568	0.00060	0.0116	
NMMR	IDW	p = 3.0	0.06475	0.04744	0.00354	−0.0649	
NMMR	Ordinary Kriging	-	0.05942	0.04316	−0.0012	0.1032	
RRD	IDW	p = 1.0	0.05457	0.04207	0.00168	−0.0442	Yes
RRD	IDW	p = 1.5	0.05604	0.04280	0.00213	−0.1014	
RRD	IDW	p = 2.0	0.05852	0.04438	0.00162	−0.2007	
RRD	IDW	p = 2.5	0.06116	0.04585	0.00079	−0.3117	
RRD	IDW	p = 3.0	0.06350	0.04730	0.00007	−0.4141	
RRD	Ordinary Kriging	-	0.05500	0.04338	0.00006	−0.0606	

For the NMMR, IDW with p = 1.5 produced the lowest RMSE (0.05732), slightly outperforming IDW with p = 1.0 (RMSE = 0.05761) and Ordinary Kriging (RMSE = 0.05942). The selected model also produced MAE = 0.04153, ME = 0.00089, and R^2 = 0.1655. Increasing the IDW power beyond 1.5 progressively reduced model performance, with RMSE increasing to 0.06475 at p = 3.0. For the RRD, IDW with p = 1.0 yielded the lowest RMSE (0.05457), followed by Ordinary Kriging (0.05500). The selected IDW model had MAE = 0.04207 and ME = 0.00168, whereas its R^2 was - 0.0442. Higher IDW powers resulted in progressively larger RMSE values, reaching 0.06350 at p = 3.0.

The near-zero ME values for both selected models indicate little systematic overall bias in the cross-validation predictions. However, the R^2 values provide an important qualification. The NMMR model achieved an R^2 of only 0.1655, while the RRD model produced a negative R^2 of - 0.0442. These values indicate that the spatial coordinates alone explained only a limited proportion of the variation in SQI, particularly in the RRD. The decline in performance as IDW power increases also suggests that assigning excessive influence to nearby observations was not advantageous for this dataset. A relatively moderate power of 1.5 performed best in the NMMR, whereas the lower power of 1.0 was optimal in the RRD. This difference indicates that the spatial structure of SQI was not identical between the two regions. Nevertheless, the relatively weak cross-validation performance suggests that the interpolated

SQI surfaces should be interpreted with caution. The maps are more appropriately considered regional-scale representations of spatial patterns than highly accurate predictions of SQI at individual unsampled locations. Factors such as soil type, land use, agricultural management, parent material, and localized PTE inputs may introduce spatial variation that cannot be captured adequately by distance-based interpolation alone (Anh et al., 2026; Kutakimina et al., 2026; Tuan Anh et al., 2026). These results indicate that the optimal interpolation approach was region-specific, highlighting the importance of empirical cross-validation when selecting spatial interpolation methods rather than assuming that a single method or parameterization is universally optimal (Li & Heap, 2014).

Spatial distribution of the Soil Quality Index

The spatial distribution of SQI generated using the selected interpolation models is shown in Fig. 5. The map reveals that SQI values across both regions were predominantly within an intermediate range, but clear spatial heterogeneity was evident, particularly in the NMMR.

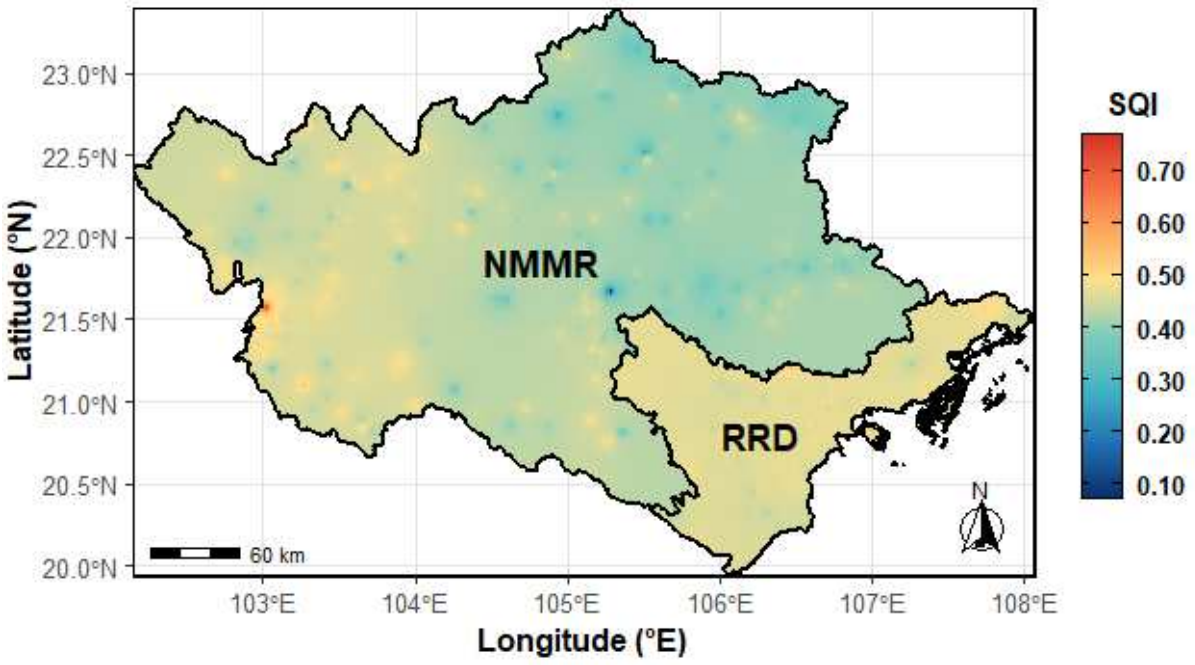


Fig. 5 Spatial distribution of the soil quality index across the NMMR and RRD regions.

The NMMR exhibited a broader spatial range of SQI, including localized areas with relatively low values of approximately 0.1 - 0.3 and areas with higher values approaching or exceeding 0.6 - 0.7. This spatial heterogeneity is consistent with the wider range of observed

SQI values reported in Table 4. In contrast, the RRD showed a comparatively narrower spatial range, with most areas characterized by SQI values around approximately 0.4 - 0.5. The spatial pattern therefore broadly agrees with the statistical comparison, which showed a higher mean SQI and a smaller overall range in the RRD. The map also demonstrates why relying solely on regional mean SQI may obscure localized areas of relatively poor soil quality. Although the NMMR had a lower mean SQI, it also contained areas with relatively high SQI. Conversely, the RRD had a higher mean SQI but did not exhibit uniformly high values across the entire region. The localized low-SQI areas in the NMMR may therefore represent priority locations for additional field investigation as reported in previous studies (Balkrishna et al., 2026; De et al., 2022; Khan et al., 2026; Soliman et al., 2026). However, given the relatively weak cross-validation performance, particularly the low R^2 values reported in Table 5, these areas should be interpreted as screening or prioritization zones rather than confirmed soil-quality hotspots. Independent sampling would be required to verify the spatial extent and magnitude of these low-SQI areas. The spatial patterns observed in Fig. 5 may reflect differences in local soil properties, land-use history, terrain, and other environmental factors (Anh et al., 2026; Clemens et al., 2010; Tuan Anh et al., 2026). However, these factors were not explicitly incorporated into the interpolation models in the present study. Consequently, causal interpretation of individual spatial patterns should be avoided without additional explanatory data.

Integrated interpretation of soil quality in the NMMR and RRD

The combined results indicate that soil quality in the NMMR and RRD is characterized by a significant regional difference superimposed on substantial within-region heterogeneity. At the individual-indicator level, the NMMR had significantly higher bulk density and Cr concentrations, whereas the RRD had significantly higher total N. Most of the other measured soil properties and PTEs did not differ significantly between the two regions. PCA reduced the multidimensional dataset to a smaller set of representative indicators, yielding an MDS comprising Pb, As, bulk density, Cd, Zn, organic C, and clay. The selected indicators captured both PTE-related variability and fundamental soil physical and chemical characteristics, consistent with the rationale of MDS-based SQI approaches that aim to retain representative information while reducing redundancy among soil indicators (Raiesi, 2017; Pujia Yu et al., 2018). Entropy weighting subsequently assigned the largest individual contribution to Pb (17.03%) and the smallest to clay (8.06%). The resulting SQI was

significantly higher in the RRD than in the NMMR, with mean values of 0.4633 and 0.4280, respectively, corresponding to an 8.25% difference and a moderate effect size (Cohen's d = 0.581). However, the wider range of SQI values in the NMMR demonstrates that its lower mean value does not imply uniformly poorer soil quality throughout the region. The spatial analysis further supports this interpretation. The NMMR displayed greater spatial heterogeneity, whereas the RRD showed a relatively more homogeneous SQI pattern. Nevertheless, the relatively low cross-validation R^2 values indicate that the spatial structure of SQI was only weakly captured by the interpolation models, highlighting the importance of validating interpolation performance before using predicted surfaces for site-specific interpretation (Li & Heap, 2014; Xu & Zhang, 2023). Taken together, the findings suggest that the PCA - MDS entropy framework provides a useful regional-scale screening approach for integrating multiple dimensions of soil quality. The resulting SQI and spatial maps can help identify areas requiring further investigation, but they should not be interpreted as definitive assessments of site-specific soil condition without additional validation.

Conclusion

An integrated PCA-MDS entropy-weighting framework was applied to assess and spatially characterize soil quality in the Northern Mountainous and Midland Region (NMMR) and the Red River Delta (RRD) of northern Vietnam. Seven indicators, including Pb, As, bulk density, Cd, Zn, organic carbon, and clay, were selected as the Minimum Data Set, representing both potentially toxic element-related variability and key soil physical and chemical properties. The mean SQI was significantly higher in the RRD (0.4633 ± 0.0537) than in the NMMR (0.4280 ± 0.0629), corresponding to an 8.25% difference and a moderate effect size (Cohen's d = 0.581). The NMMR also exhibited greater variability in SQI, indicating substantial within-region heterogeneity rather than uniformly lower soil quality. Spatial interpolation further revealed heterogeneous SQI patterns, particularly in the NMMR; however, the relatively low cross-validation performance indicates that the predicted surfaces should be interpreted primarily as regional-scale screening tools rather than precise site-specific estimates. The PCA-MDS entropy framework highlights a practical approach to

integrating multiple dimensions of soil quality and identifying areas warranting further investigation and targeted soil management in northern Vietnam.

Funding

This research was funded by Insitute of Earth Science - Viet Nam Academy of Science and Technology under project number NVKN 13.02/26-27.

Author Contributions

Duong Thi Lim: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing - original draft; Dang Tran Quan, Nguyen Thi Hue, and Nguyen Thi Lan Huong: Investigation, Data curation; Le Thi Phuong Dung, Nguyen Thi Huong Thuy, and Nguyen Thi Thu: Investigation, Data curation; Cu Sy Thang and Nguyen Viet Cuong: Methodology, Investigation; Nguyen Thi Thu Trang and Nguyen Duc Hung: Investigation, Data curation; Nguyen Ngoc Quy: Methodology; Nguyen-Sy Toan: Conceptualization, Methodology, Visualization, Writing - review & editing, Supervision. All authors contributed to the interpretation of the results, reviewed and approved the final version of the manuscript.

Conflict of interest

The authors declare that they have no known competing fnancial interests or personal relationships that could have appeared to infuence the work reported in this paper.

Statements and Declarations

All authors have read, understood, and have complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors.

References

Abdul-Wahab, D., Sunkari, E. D., Yalley, C. A., Asare, E. A., Laar, C., Chakrabarti, T.,…Ambushe, A. A. (2026). Competing geochemical drivers and probabilistic health risk assessment of groundwater in the semi-arid Voltaian sedimentary basin, Ghana. *Sci Rep*. https://doi.org/10.1038/s41598-026-66475-4

Almasoud, F. I., Usman, A. R., & Al-Farraj, A. S. (2014). Heavy metals in the soils of the Arabian Gulf coast affected by industrial activities: analysis and assessment using enrichment factor and multivariate analysis. *Arabian Journal of Geosciences*, *8*(3), 1691-1703. https://doi.org/10.1007/s12517-014-1298-x

Anh, T. T., Lim, D. T., Dang, P. T., Ha, N. M., Huong, N. T. L., Quan, D. T.,…Toan, N. S. (2026). Spatial variability and environmental controls of trace elements (Co, Cu, Mo, Ni, and Zn) in topsoils of Cao Bang Province, Northern Vietnam. *Environ Monit Assess*, *198*(7). https://doi.org/10.1007/s10661-026-15588-4

Balkrishna, A., Ghosh, S., Banerjee, S., Arya, V., Kaushik, I., Semwal, D.,…Pillai, S. U. (2026). Evaluation of region-specific organic farming suitability in the Indo-Gangetic basin using the geographical information system based soil quality index and machine learning approach. *Discover Soil*, *3*(1). https://doi.org/10.1007/s44378-026-00211-2

Borchers, A., & Pieler, T. (2010). Programming pluripotent precursor cells derived from Xenopus embryos to generate specific tissues and organs. *Genes (Basel)*, *1*(3), 413-426. https://doi.org/10.3390/genes1030413

Cai, L. M., Wang, Q. S., Luo, J., Chen, L. G., Zhu, R. L., Wang, S., & Tang, C. H. (2019). Heavy metal contamination and health risk assessment for children near a large Cu-smelter in central China. *Sci Total Environ*, *650*(Pt 1), 725-733. https://doi.org/10.1016/j.scitotenv.2018.09.081

Clemens, G., Fiedler, S., Cong, N. D., Van Dung, N., Schuler, U., & Stahr, K. (2010). Soil fertility affected by land use history, relief position, and parent material under a tropical climate in NW-Vietnam. *CATENA*, *81*(2), 87-96. https://doi.org/10.1016/j.catena.2010.01.006

De, P., Deb, S., Deb, D., Chakraborty, S., Santra, P., Dutta, P.,…Choudhury, A. (2022). Soil quality under different land uses in eastern India: Evaluation by using soil indicators and quality index. *PLoS One*, *17*(9), e0275062. https://doi.org/10.1371/journal.pone.0275062

Ghosh, T., Das, B., Chakraborty, D., Singh, V. K., Das, D., Aggarwal, P.,…Prasad, P. V. V. (2025). Evaluation of soil quality frameworks in rice-wheat systems under integrated nutrient management in Indo-Gangetic plains. *Sci Rep*, *15*(1), 37511. https://doi.org/10.1038/s41598-025-21630-1

Guo, F., Zhang, A., Liu, X., Yuan, S., Chang, Z., Zhu, L.,…Xiao, X. (2026). Construction of a comprehensive bio-toxicity index for heavy metals across multiple trophic levels based on the integration of lethal concentration-response curves using entropy weighting method. *Frontiers in Environmental Science*, *14*. https://doi.org/10.3389/fenvs.2026.1836960

Jahany, M., & Rezapour, S. (2020). Assessment of the quality indices of soils irrigated with treated wastewater in a calcareous semi-arid environment. *Ecological Indicators*, *109*, 105800. https://doi.org/10.1016/j.ecolind.2019.105800

Khan, S. A., Khan, A. R., & Said, S. (2026). Performance evaluation of soil quality indices in agricultural landscapes. *Discover Soil*, *3*(1). https://doi.org/10.1007/s44378-026-00267-0

Kutakimina, P. B., Mbanga, J.-A. L., Mulaji, C. K., Carvalho, F. P., Poté, J., & Atibu, E. K. (2026). Geochemical evaluation of trace metals and rare-earth elements in soils and sediments from a legacy copper mining site in Kongo Central, Democratic Republic of the Congo. *Discover Geoscience*, *4*(1). https://doi.org/10.1007/s44288-026-00428-2

Li, J., & Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, *53*, 173-189. https://doi.org/10.1016/j.envsoft.2013.12.008

Luo, Z., Feng, W., Luo, Y., Baldock, J., & Wang, E. (2017). Soil organic carbon dynamics jointly controlled by climate, carbon inputs, soil properties and soil carbon fractions. *Glob Chang Biol*, *23*(10), 4430-4439. https://doi.org/10.1111/gcb.13767

Maaz, T. M., Heck, R. H., Glazer, C. T., Loo, M. K., Zayas, J. R., Krenz, A.,…Deenik, J. L. (2023). Measuring the immeasurable: A structural equation modeling approach to assessing soil health. *Sci Total Environ*, *870*, 161900. https://doi.org/10.1016/j.scitotenv.2023.161900

Mamehpour, N., Rezapour, S., & Ghaemian, N. (2021). Quantitative assessment of soil quality indices for urban croplands in a calcareous semi-arid ecosystem. *Geoderma*, *382*, 114781. https://doi.org/10.1016/j.geoderma.2020.114781

Mussgnug, F., Becker, M., Son, T. T., Buresh, R. J., & Vlek, P. L. G. (2006). Yield gaps and nutrient balances in intensive, rice-based cropping systems on degraded soils in the Red

River Delta of Vietnam. *Field Crops Research*, *98*(2-3), 127-140. https://doi.org/10.1016/j.fcr.2005.12.012

Nguyen-Sy, T., Nguyen Thi, L. P., Duong Thi, L., Do Hong, H., Tran Thi, P., Huynh, H.,…Phan Chi, U. (2025). Spatial variation in soil C, N, and trace elements in paddy soil across the coastal, inland, and upland regions of central Vietnam. *Environ Monit Assess*, *197*(9), 977. https://doi.org/10.1007/s10661-025-14422-7

Nguyen-Sy, T., Tang, S., Phung, L. D., Doi, R., Shiono, H., Tawaraya, K.,…Cheng, W. (2025). Long-term changes in macronutrients, micronutrients, and potentially toxic elements under 33 years of continuous chemical fertilizer and organic matter application. *Soil Science and Plant Nutrition*, *71*(6), 758-767. https://doi.org/10.1080/00380768.2025.2558604

Pei, J., Li, J., Luo, Y., Rillig, M. C., Smith, P., Gao, W.,…Nie, M. (2025). Patterns and drivers of soil microbial carbon use efficiency across soil depths in forest ecosystems. *Nat Commun*, *16*(1), 5218. https://doi.org/10.1038/s41467-025-60594-8

Phuong, N. M., Kang, Y., Sakurai, K., Iwasaki, K., Kien, C. N., Van Noi, N., & Son, L. T. (2008). Arsenic contents and physicochemical properties of agricultural soils from the Red River Delta, Vietnam. *Soil Science and Plant Nutrition*, *54*(6), 846-855. https://doi.org/10.1111/j.1747-0765.2008.00312.x

Phuong, N. M., Kang, Y., Sakurai, K., Iwasaki, K., Kien, C. N., Van Noi, N., & Son, L. T. (2009). Levels and Chemical Forms of Heavy Metals in Soils from Red River Delta, Vietnam. *Water, Air, and Soil Pollution*, *207*(1-4), 319-332. https://doi.org/10.1007/s11270-009-0139-0

Rahmanipour, F., Marzaioli, R., Bahrami, H. A., Fereidouni, Z., & Bandarabadi, S. R. (2014). Assessment of soil quality indices in agricultural lands of Qazvin Province, Iran. *Ecological Indicators*, *40*, 19-26. https://doi.org/10.1016/j.ecolind.2013.12.003

Raiesi, F. (2017). A minimum data set and soil quality index to quantify the effect of land use conversion on soil quality and degradation in native rangelands of upland arid and semiarid regions. *Ecological Indicators*, *75*, 307-320. https://doi.org/10.1016/j.ecolind.2016.12.049

Soliman, A. M., Elbshbeshi, A., Belal, A. B., Amira, M. S., Hassan, H. A., Abdelhameed, H. H., & Ibraheem, I. M. (2026). Integrated GIS-based Assessment of Soil Quality Using the MEDALUS Framework in the Nile Delta, Egypt. *Earth Systems and Environment*, *10*(6), 7169-7185. https://doi.org/10.1007/s41748-025-01005-2

Stenberg, B. (1999). Monitoring Soil Quality of Arable Land: Microbiological Indicators. *Acta Agriculturae Scandinavica, Section B - Soil & Plant Science*, *49*(1), 1-24. https://doi.org/10.1080/09064719950135669

Sultanovna, M. G., & Ismailovich, R. M. (2026). Integrated Assessment of Soil Quality and Heavy Metal Contamination in Arid Agricultural Soils of Northwestern Uzbekistan. *American Journal Of Agriculture And Horticulture Innovations*, *06*(04), 16-22. https://doi.org/10.37547/ajahi/volume06issue03-03

Toan, N. S., Ngoc-Son, T., Van An, N., Hanh, D. H., Phu, T. T., Hai, H.,…Uyen, P. C. (2025). Impact of land-use conversion from a long-term paddy to abandoned field on the distribution of C, N, and trace elements: a study in central Viet Nam. *Environ Geochem Health*, *47*(12), 559. https://doi.org/10.1007/s10653-025-02886-y

Tuan Anh, T., Thi Lim, D., Thanh Dang, P., Manh Ha, N., Lan Huong, N. T., Tran Quan, D.,…Toan, N.-S. (2026). Influence of topography, elevation, and parent materials on the distribution of potentially toxic elements (PTEs) in Surface soils: a case study in northern vietnam. *Soil and Sediment Contamination: An International Journal*, 1-18. https://doi.org/10.1080/15320383.2025.2612286

Uymaz, S. E., Bülbül, U., Özakın, B., & Gültekin, K. (2026). Entropy-based integrated multi-criteria decision-making framework for selecting vegetable lubricants in manufacturing processes. *Engineering Research Express*, *8*(15), 155536. https://doi.org/10.1088/2631-8695/ae8e33

Wang, C., Hu, J., Zhang, Y., Di, Y., & Wu, X. (2025). Spatial distribution characteristic, source apportionment, and risk assessment of heavy metals in the soil of an urban riparian zone. *Ecotoxicol Environ Saf*, *298*, 118271. https://doi.org/10.1016/j.ecoenv.2025.118271

Xie, T., Wang, M., Su, C., & Chen, W. (2018). Evaluation of the natural attenuation capacity of urban residential soils with ecosystem-service performance index (EPX) and entropy-weight methods. *Environ Pollut*, *238*, 222-229. https://doi.org/10.1016/j.envpol.2018.03.013

Xu, H., & Zhang, C. (2023). Development and applications of GIS-based spatial analysis in environmental geochemistry in the big data era. *Environ Geochem Health*, *45*(4), 1079-1090. https://doi.org/10.1007/s10653-021-01183-8

Yang, W., Yue, J., Liu, Q., Liang, S., & Ye, Y. (2026). Soil quality assessment based on the TDS and MDS methods in the Qinling mountains of China. *PLoS One*, *21*(8), e0335384. https://doi.org/10.1371/journal.pone.0335384

Yeilagi, S., Rezapour, S., & Asadzadeh, F. (2021). Degradation of soil quality by the waste leachate in a Mediterranean semi-arid ecosystem. *Sci Rep*, *11*(1). https://doi.org/10.1038/s41598-021-90699-1

Yu, P., Liu, S., Zhang, L., Li, Q., & Zhou, D. (2018). Selecting the minimum data set and quantitative soil quality indexing of alkaline soils under different land uses in northeastern China. *Science of The Total Environment*, *616-617*, 564-571. https://doi.org/10.1016/j.scitotenv.2017.10.301

Yu, P., Liu, S., Zhang, L., Li, Q., & Zhou, D. (2018). Selecting the minimum data set and quantitative soil quality indexing of alkaline soils under different land uses in northeastern China. *Sci Total Environ*, *616-617*, 564-571. https://doi.org/10.1016/j.scitotenv.2017.10.301

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Fig.S1.docx