International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Hybrid Quantum–Classical Machine Learning Framework for Village-Level Soil Environmental Sustainability and Land Degradation Risk Assessment in Bilaspur District, India

Chandrashekhar[1], Amit Kumar Chandanan[2]

[1]Ph.D. Research Scholar, Department of Computer Science and Information Technology, Guru Ghasidas Vishwavidyalaya Bilaspur, India
[2]Associate Professor, Department of Computer Science and Engineering, Guru Ghasidas Vishwavidyalaya Bilaspur, India

Abstract

The presence of multivariate dependency relationships, spatial heterogeneity, and nonlinear nutrient interactions makes it challenging to evaluate the environmental sustainability of soils accurately using conventional statistical models. For the Bilaspur district in the state of Chhattisgarh, India, this research proposes a hybrid quantum-classical machine learning approach for land degradation risk classification and soil environmental sustainability.

A structured dataset of 677 soil samples from 17 villages spread across four administrative blocks was analyzed using twelve micronutrient and physicochemical parameters (pH, EC, OC, N, P, K, S, Zn, Fe, Mn, Cu, and B). To translate standardized nutrient variables into understandable sustainability descriptors, composite environmental indicators such as the Soil Nutrient Balance Index (SNBI), Soil Fertility Index (SFI), Micronutrient Deficiency Index (MDI), Soil Degradation Index (SDI), and Environmental Sustainability Index (ESI) were created mathematically. After removing scale dominance effects with Min–Max normalization, deep autoencoder-based nonlinear dimensional compression (5→16→8→4 architecture) was performed using Mean Squared Error optimization and ReLU activation.

After being extracted, the 4-dimensional latent representation was embedded into a 4-qubit ZZFeatureMap and rescaled to the angular domain [0, π]. Sustainability risk classification was carried out by a precomputed Quantum Support Vector Machine (QSVM) using ESI quantile thresholds, and quantum state overlap was calculated using a fidelity-based Quantum Kernel Estimation (QKE) framework. With a balanced distribution of classes, the suggested model's accuracy was 88.23%.

The findings show that quantum kernel learning successfully captures nonlinear environmental sustainability structures, offering a scalable decision-support tool for sustainable land-use planning, region-specific soil management, and precision agriculture.

Keywords: Soil Environmental Sustainability, Quantum Machine Learning (QML), Quantum Support Vector Machine (QSVM), Precision Agriculture, Soil Nutrient Imbalance, Environmental Sustainability

Index (ESI), Bilaspur District Soil Study.

1. Introduction

Soil nutrient disorder, micronutrient deficiencies, and the resulting land degradation are increasingly becoming a threat to sustainable agricultural productivity, especially in region-specific agro-ecosystems where environmental stress is still largely unreported at the village level. Long-term agricultural productivity and ecological sustainability are significantly affected by the deterioration of soil health due to disorders in nitrogen (N), phosphorus (P), potassium (K), organic carbon (OC), pH extremes, electrical conductivity (EC), and micronutrient deficiencies such as zinc (Zn) and boron (B) [1], [2]. The planning for effective environmental sustainability is hampered by the fact that traditional soil fertility estimation techniques are incapable of comprehending complex interdependencies between physicochemical and micronutrient variables [3].

Using data-driven approaches, recent advances in machine learning (ML) have demonstrated potential in estimating soil fertility and identifying the risk of land degradation [4], [5]. However, conventional machine learning algorithms may become less effective for nonlinear relationships between multiple soil variables, particularly in a heterogeneous agricultural landscape [6]. Autoencoders and other deep learning-based dimensionality reduction techniques have been used to identify latent soil variables in order to enhance classification performance [7]. Furthermore, quantum machine learning (QML) models, specifically Quantum Support Vector Machines (QSVM), have shown enhanced capacity to capture high-dimensional environmental relationships through the use of quantum feature mapping and entanglement-based kernels [8], [9].

In order to develop an ICAR-based QML approach for evaluating environmental sustainability, the current study uses twelve macro, micro, and physicochemical parameters of soil collected from seventeen villages located over five blocks in the Bilaspur district [10]. Several environmental indices, including the Soil Nutrient Balance Index (SNBI), Soil Fertility Index (SFI), Micronutrient Deficiency Index (MDI), Soil Degradation Index (SDI), and Environmental Sustainability Index (ESI), were developed to evaluate the effect of nutrients on environmental stress at the village level. The risk zones for soil sustainability were identified using quantum angle encoding and QSVM classification, while a deep autoencoder was employed for feature compression.

The proposed approach is expected to facilitate region-specific soil management practices, sustainable fertilizer use, and land use policy planning for soil sustainability [11,12]. Future work can be done by incorporating climatic factors and crop productivity measures to improve the predictive assessment of environmental impact for long-term agro-ecosystem sustainability.

2. Literature Review

Soil nutrient imbalance and progressive land degradation are major obstacles to long-term agricultural production and environmental sustainability, especially in region-specific agro-ecosystems where localized soil degradation patterns are frequently undiagnosed. Soil health indices such as nitrogen (N), phosphorus (P), potassium (K), organic carbon (OC), pH, electrical conductivity (EC), and micronutrients are critical for nutrient cycling, ecosystem resilience, and food security [1,2]. In the Indian scenario, large-scale studies have revealed the prevalence of micronutrient deficiencies and soil alkalinity, thereby underlining the pressing need for soil sustainability monitoring [14]. In addition, unbalanced fertilizer

practices and conservation tillage systems have been observed to affect soil productivity and environmental sustainability [3], thereby reiterating the need for data assessment frameworks.

Machine learning (ML) methods have been increasingly used in agricultural and environmental applications to model complex interactions between climatic, biological, and soil variables. Various studies have shown the high predictive accuracy of crop yield forecasting using ensemble methods such as Random Forest and XGBoost [4] Similarly, hyperspectral imaging using UAVs and support vector machines have been shown to be effective in estimating the variability of soil organic matter and nutrients [7,11]. However, traditional ML models have been found to be less effective in modeling complex nonlinear interactions between various soil nutrient variables, thus being less applicable in the classification of sustainability risks.

Recent breakthroughs in quantum machine learning (QML) have offered improved computational power by representing classical inputs as high-dimensional Hilbert spaces through quantum feature maps and parameterized quantum circuits [9,10]. Quantum neural networks have been shown to possess supervised learning abilities on structured data inputs through near-term quantum computing [11]. Moreover, quantum kernel-based models have been shown to possess competitive or better classification accuracy than classical SVM methods in multiclass real-world scenarios [13,14]. Benchmarking studies also emphasize the significance of fidelity-based kernel representation in enhancing generalization performance [15,16], these developments indicate the potential of integrating soil nutrient indicators with QML frameworks for region-specific environmental sustainability assessment and village-level land degradation risk mapping, particularly in heterogeneous agricultural landscapes such as Bilaspur district.

2.1 Literature Comparison

Table 1: Literature Comparison Table

Study	Model Used	Data Type	Environmental Focus	Learning Capability	Quantum Advantage	Applicability
[13]	Statistical	Soil Survey	Soil Health	Low	No	Policy
[4]	RF, XGB	Agro-data	Yield	Moderate	No	Crop
[14]	SVM, ELM	UAV Soil Data	Nutrient Mapping	Moderate	No	Precision Agri
[10]	PQC	Classical	Pattern Learning	High	Yes	ML Tasks
[11]	QNN	Classical	Classification	High	Yes	Structured Data
[15]	QSVM	Real Dataset	Pattern Classification	High	Yes	Multiclass
[16]	FQK/PQK	Structured	Regression	High	Yes	Kernel Tasks
Proposed Work	QSVM + AE	Soil Nutrients	Sustainability Risk	Very High	Yes	Village-level

2.2 Problem Statement

At localized scales, soil nutrient imbalance, micronutrient deficiency, and pH variability have a major impact on long-term agricultural productivity and environmental sustainability. Due to a lack of predictive analytical frameworks, village-level land degradation patterns frequently go unnoticed in areas like Bilaspur district.

Conventional soil assessment techniques fail to model nonlinear nutrient interactions affecting sustainable land-use planning.

2.3 Research Gap

Most agro-environmental research is centered on crop productivity or climate-related variables and does not often combine multi-nutrient soil variables for sustainability analysis.

Conventional machine learning algorithms are not suited to handle the complex physicochemical relationships that exist in soil data.

There is a lack of studies on village-level environmental risk classification, especially using quantum machine learning for sustainable monitoring systems.

2.4 Limitations of existing studies

Most existing methods for assessing the sustainability of soils in terms of environmental stress due to nutrient imbalance are not predictive. Machine learning algorithms are mostly concerned with yield prediction and not soil degradation analysis, and remote sensing-based methods are mostly crop-level. Conventional computational models are not suited for high-dimensional soil data with nonlinear interactions.

2.5 Research Motivation

The increasing rate of soil degradation and micronutrient depletion requires a region-specific predictive model for sustainability assessment at the village level. Combining soil nutrient indicators with quantum-assisted machine learning models such as QSVM can improve the classification of risk zones. This helps in soil management, fertilizer application, and land use planning for sustainable soil conditions.

3. Material and Methods

3.1 Materials

The study area for this paper is located in the Bilaspur district of Chhattisgarh state, which falls in the agro-ecological region of 21.47° to 23.08° N latitude and 81.14° to 83.15° E longitude. A total of 677 soil samples were collected from 17 villages in four blocks: Bilha, Masturi, Takhatpur, and Bilaspur. There are 12 key soil parameters in this dataset: pH, EC, OC, N, P, K, S, Zn, Fe, Mn, Cu, and B. Village and block variables were encoded for structured computational modeling for spatial sustainability assessment.

3.2 Methodology

The suggested framework uses a hybrid classical-quantum machine learning model based on the soil physicochemical and micronutrient parameters of seventeen villages in five administrative blocks of the Bilaspur district to evaluate the risk of land degradation and the sustainability of the soil at the village level.

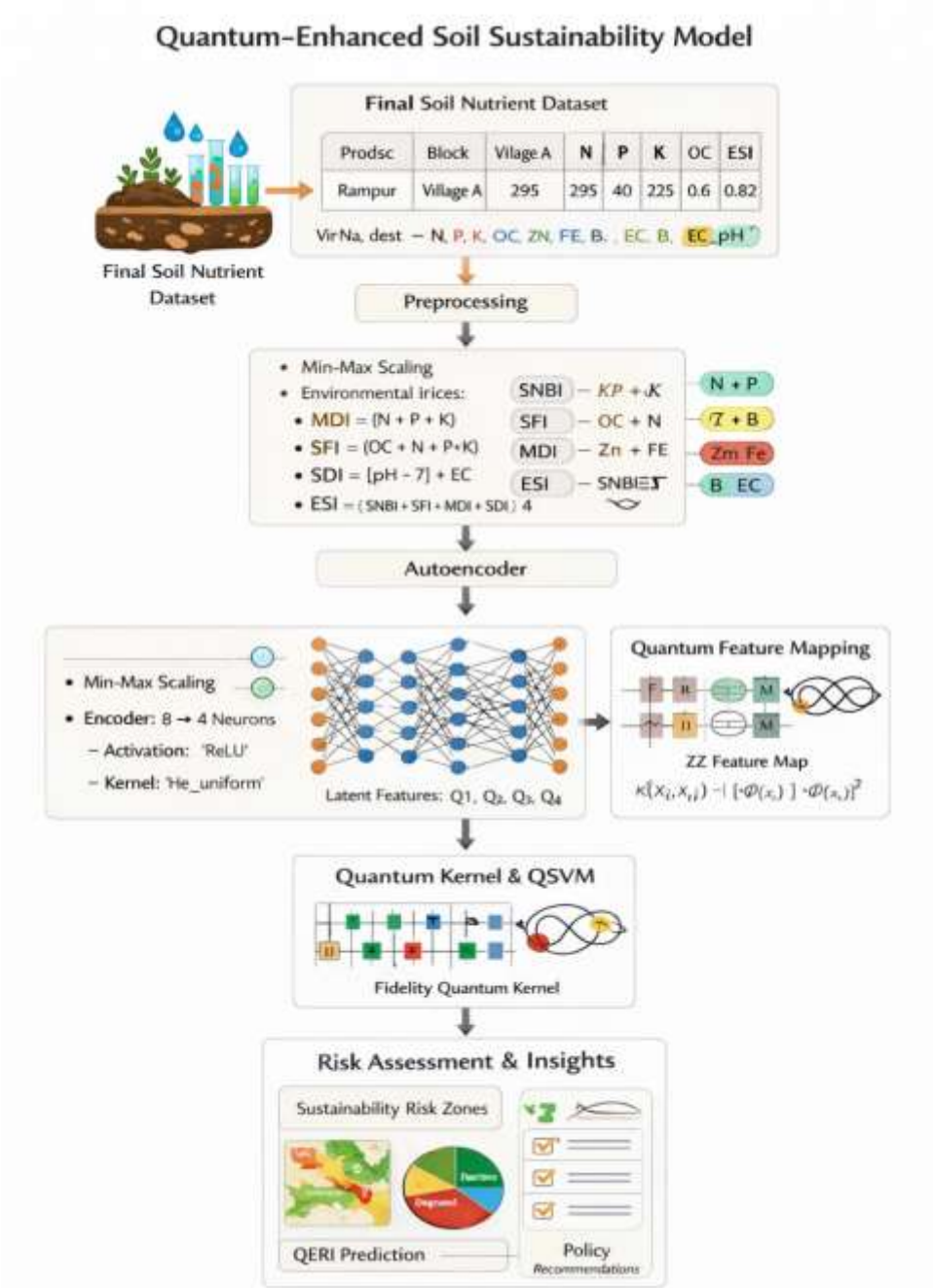

Figure 1: Proposed QML-based framework for village-level soil environmental sustainability and land degradation risk assessment in Bilaspur district.

3.2.1 Data Preprocessing

Before developing the model, systematic preprocessing techniques were used to guarantee data reliability and model robustness.

Step 1: Acquisition of the Original Dataset

The dataset includes 677 observations and environmental soil parameters gathered from four administrative blocks and 17 villages.

Raw attributes : {PH, EC, OC, N, P, K, S, B, ZN, FE, MN, CU, VILLAGE, Block}

Total raw attributes = 14

Step 2: Categorical Encoding

To facilitate computer modeling, categorical variables were converted into numerical representations:

$$VILLAGE \rightarrow Village\_Encoded$$
$$Block \rightarrow Block\_Encoded$$

Encoding function:

$$E(x) = \text{LabelEncoding}(x)$$

After encoding: {PH, EC, OC, N, P, K, S, B, ZN, FE, MN, CU, Village, Block, Village_Encoded, Block_Encoded}

Total attributes after encoding = 16

Step 3: Preliminary Data Cleaning

Since the dataset was preprocessed beforehand:

$$\sum I(x_i = NaN) = 0$$
$$D_{clean} = D$$
$$D_{unique} = D$$

There were no missing or duplicate values discovered.

As a result, no further imputation or filtering was necessary at this point.

Step 4: Nutrient Feature Selection

Although 12 environmental parameters are available: {N, P, K, S, B, ZN, FE, MN, CU, PH, EC, OC}

For sustainability index formulation and QML modelling, selected features were:

$$X = \{N, P, K, B, ZN, FE, PH, EC, OC\}$$

Expanded vector:

$$X_i = (N_i, P_i, K_i, B_i, ZN_i, FE_i, PH_i, EC_i, OC_i)$$

Where $i = 1, \ldots, 677$

Step 5: Statistical Standardization (Already Applied in Dataset)

Observed dataset statistics:

$$\mu \approx 0, \sigma \approx 1$$

This confirms prior Z-score normalization:

$$X_{standardized} = \frac{X-\mu}{\sigma} \qquad (1)$$

Thus, no additional standardization was required.

Step 6: Feature Scaling (Min-Max Normalization)

To ensure uniform contribution of each parameter, Min-Max scaling was applied:

$$X_{scaled} = \frac{X-X_{min}}{X_{max}-X_{min}} \qquad (2)$$

Where: X= observed value, X_{min}= minimum dataset value, X_{max}= maximum dataset value

This adjustment normalizes the feature space between 0 and 1, decreasing scale-induced bias during model learning.

Nutrient-wise Computation (Using Dataset Means)

Nitrogen (N)

$$N_{scaled} = \frac{0.0268 - (-2.8115)}{2.3306 - (-2.8115)} = \frac{2.8383}{5.1421} = 0.552$$

Phosphorus (P)

$$P_{scaled} = \frac{-0.1579 - (-3.8917)}{2.4473 - (-3.8917)} = \frac{3.7338}{6.3390} = 0.589$$

Potassium (K)

$$K_{scaled} = \frac{-0.2105 - (-2.9200)}{2.0759 - (-2.9200)} = \frac{2.7095}{4.9959} = 0.542$$

Zinc (Zn)

$$Zn_{scaled} = \frac{-0.0075 - (-2.5612)}{2.5606 - (-2.5612)} = \frac{2.5537}{5.1218} = 0.498$$

Iron (Fe)

$$Fe_{scaled} = \frac{0.0221 - (-3.2155)}{2.6863 - (-3.2155)} = \frac{3.2376}{5.9018} = 0.548$$

Boron (B)

$$B_{scaled} = \frac{-0.0925 - (-2.2557)}{1.6468 - (-2.2557)} = \frac{2.1632}{3.9025} = 0.554$$

pH

$$pH_{scaled} = \frac{-0.1320 - (-2.2768)}{2.0257 - (-2.2768)} = \frac{2.1448}{4.3025} = 0.498$$

Electrical Conductivity (EC)

$$EC_{scaled} = \frac{-0.0761 - (-1.8726)}{2.3636 - (-1.8726)} = \frac{1.7965}{4.2362} = 0.424$$

Organic Carbon (OC)

$$OC_{scaled} = \frac{0.0203 - (-1.7221)}{2.1871 - (-1.7221)} = \frac{1.7424}{3.9092} = 0.446$$

Scaled Feature Vector

$$X_{scaled} = (N, P, K, ZN, FE, B, PH, EC, OC)$$

$$X_{scaled} = (0.552, 0.589, 0.542, 0.498, 0.548, 0.554, \quad 0.498, 0.454, 0.446))$$

3.2.2 Statistical Analysis

Step 1: Dataset Dimensionality

Given dataset:

$$D \in \mathbb{R}^{677\times14}$$

Where: $n = 677$observations, $p = 14$ total attributes, 17 villages, 4 administrative blocks, Bilaspur district Chhattisgarh.

Step 2: Mean Calculation with Actual Values

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ (3)

Where:

$\bar{X}$= sample mean (average parameter value), n = total number of observations (in this case $n = 677$), i=index of observation (i = 1,2,3,...,677), X_i= observed parameter value at ith position, and ∑=summation operator.

Implemented Mean Formula (All Nutrients)

$$\bar{N} = \frac{1}{677}\sum_{i=1}^{677} N_i = 0.0268$$

$$\bar{P} = \frac{1}{677}\sum_{i=1}^{677} P_i = -0.1579$$

$$\bar{K} = \frac{1}{677}\sum_{i=1}^{677} K_i = -0.2105$$

$$\bar{Zn} = \frac{1}{677}\sum_{i=1}^{677} Z\,n_i = -0.0075$$

$$\bar{Fe} = \frac{1}{677}\sum_{i=1}^{677} F\,e_i = 0.0221$$

$$\bar{B} = \frac{1}{677}\sum_{i=1}^{677} B_i = -0.0925$$

$$\bar{pH} = \frac{1}{677}\sum_{i=1}^{677} p\,H_i = -0.1320$$

$$\bar{EC} = \frac{1}{677}\sum_{i=1}^{677} E\,C_i = -0.0761$$

$$\bar{OC} = \frac{1}{677}\sum_{i=1}^{677} O\,C_i = 0.0203$$

The near-zero mean values of all soil nutrients serve as a foundation for earlier data standardization, resulting in statistical stability, minimum scale bias, and a solid numerical foundation for quantum-based feature encoding and sustainability classification.

Step 3: Coefficient of Variation

$$CV = \frac{\sigma}{\mu} \times 100 \qquad (4)$$

To evaluate the relative dispersion of soil nutrients before standardization, the CV was computed using σ (standard deviation) and μ (mean).

Step 4: Block-wise Mean Expansion

$$\bar{X}_{block} = \frac{1}{n_b}\sum_{i=1}^{n_b} X_i \qquad (5)$$

Now substitute actual block means:

Bilaspur Block

$$\bar{N}_{Bilaspur} = -0.1708$$
$$\bar{P}_{Bilaspur} = 0.0962$$
$$\bar{K}_{Bilaspur} = 0.0169$$
$$\bar{pH}_{Bilaspur} = -0.5708$$

Masturi Block

$$\bar{N}_{Masturi} = 0.1947$$
$$\bar{K}_{Masturi} = -0.5511$$
$$\bar{Zn}_{Masturi} = 0.3177$$

Takhatpur Block

$$\bar{EC}_{Takhatpur} = 0.1101$$
$$\bar{B}_{Takhatpur} = 0.0459$$

This substitution clearly shows inter-block nutrient heterogeneity.

3.2.3 Environmental Indicator Computation

To transform standardized soil nutrient data into structured sustainability descriptors, composite environmental indices were generated using aggregated formulations.

Soil Nutrient Balance Index (SNBI)

The Soil Nutrient Balance Index (SNBI) was defined as

$$SNBI = (N + P + K)/3 \quad (6)$$

Substituting the standardized mean values $N = 0.0268$, $P = -0.1579$, and $K = -0.2105$, we obtain

$$SNBI = (0.0268 - 0.1579 - 0.2105)/3 = -0.1139$$

This score shows the stability of nutritional balance at the village level and assesses macronutrient equilibrium.

Soil Fertility Index (SFI)

The Soil Fertility Index (SFI) was computed as

$$SFI = (OC + N + P + K)/4 \quad (7)$$

Using $OC = 0.0203$, the calculated value becomes

$$SFI = (0.0203 + 0.0268 - 0.1579 - 0.210 = -0.0803$$

This formulation combines organic carbon effect with macronutrients to provide a comprehensive fertility strength indication.

Micronutrient Deficiency Index (MDI)

The Micronutrient Deficiency Index (MDI) was defined as

$$MDI = (ZN + FE + B)/3 \quad (8)$$

and substituting $ZN = -0.0075$, $FE = 0.0221$ and $B = -0.0925$, yields

$$MDI = (-0.0075 + 0.0221 - 0.0925)/3 = -0.0260$$

This index assesses micronutrient variability and probable insufficiency trends across areas.

Soil Degradation Index (SDI)

Soil degradation intensity was quantified using

$$SDI = | PH - 7 | + EC \quad (9)$$

where $PH = -0.1320$and $EC = -0.0761$, resulting in

$$SDI = | -0.1320 - 7 | -0.0761 = 7.0559$$

This formulation emphasizes deviation from neutral pH and conductivity influence, thereby representing physicochemical stress.

Environmental Sustainability Index (ESI)

Finally, the Environmental Sustainability Index (ESI) was computed as

$$ESI = (SNBI + SFI + MDI + SDI)/4 \quad (10)$$

Producing

$$ESI = (-0.1139 - 0.0803 - 0.0260 + 7.0559)/4 = 1.7089$$

The ESI combines degradation intensity, micronutrient dynamics, fertility status, and dietary balance into a single sustainability descriptor.

The resulting indices facilitate the conversion of standardized nutritional data into significant environmental concepts, resulting in a structured feature space that is then utilized for feature scaling, latent neural mapping, and quantum concepts-based sustainability risk categorization.

3.2.4 Feature Scaling Using Min–Max Normalization

Once the five environmental sustainability indicators have been calculated

$$X = \{SNBI, SFI, MDI, SDI, ESI\}$$

The structured feature matrix used for modelling was developed by systematically extracting these values from the data frame. However, the indicators showed different scales of values (SNBI ≈ −0.1139, SFI ≈ −0.0803, MDI ≈ −0.0260, SDI ≈ 7.0559, ESI ≈ 1.7089), with the SDI having a much larger scale. This could potentially cause an unbalanced training of the neurons, with the SDI dominating the feature learning process.

To ensure uniform contribution of each environmental indicator, Min–Max normalization was applied:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (11)$$

Where:

X= original environmental indicator value, X_{min}= minimum value of that indicator across all 677 samples, X_{max}= maximum value of that indicator across all samples, $X_{scaled} \in [0,1]$.

For each feature separately:

SNBI Scaling

$$SNBI_{scaled} = \frac{SNBI - SNBI_{min}}{SNBI_{max} - SNBI_{min}}$$

SFI Scaling

$$SFI_{scaled} = \frac{SFI - SFI_{min}}{SFI_{max} - SFI_{min}}$$

MDI Scaling

$$MDI_{scaled} = \frac{MDI - MDI_{min}}{MDI_{max} - MDI_{min}}$$

SDI Scaling

$$SDI_{scaled} = \frac{SDI - SDI_{min}}{SDI_{max} - SDI_{min}}$$

ESI Scaling

$$ESI_{scaled} = \frac{ESI - ESI_{min}}{ESI_{max} - ESI_{min}}$$

After applying Min–Max normalization, the transformed indicators satisfied $X_{scaled}^{min} = 0$ and $X_{scaled}^{max} = 1$ for each feature. Consequently, for all 677 samples,

$$0 \leq SNBI_{scaled}, SFI_{scaled}, MDI_{scaled}, SDI_{scaled}, ESI_{scaled} \leq 1$$

The resulting scaled matrix $X_{scaled} \in \mathbb{R}^{677\times5}$ forms the normalized feature space, expressed as

$$X_{scaled} = (SNBI_s, SFI_s, MDI_s, SDI_s, ESI_s)$$

This bounded and dimensionally consistent representation serves as the direct input to the autoencoder neural architecture for subsequent latent feature extraction.

3.2.5 Autoencoder-Based Latent Feature Extraction

A deep autoencoder network was used to the normalized environmental feature matrix $X_{scaled} \in \mathbb{R}^{677\times5}$ to capture nonlinear sustainability relationships while reducing feature dimensionality. The encoder transformation goes as follows:

$$Z = f(WX_{scaled} + b) \quad (12)$$

where X_{scaled}represents the Min–Max normalized indicators $(SNBI_s, SFI_s, MDI_s, SDI_s, ESI_s)$, Wdenotes the learnable weight matrices, $b$the bias vectors, $f(\cdot)$the ReLU activation function, and $Z$the latent sustainability representation.

The encoder design comprised of a 5-dimensional input (scaled indicators), two hidden layers of 16 and 8 neurons with ReLU activation, and a 4-neuron latent layer. So, the compression mapping was:

$$\mathbb{R}^5 \rightarrow \mathbb{R}^{16} \rightarrow \mathbb{R}^8 \rightarrow \mathbb{R}^4$$

The resulting latent representation is:

$$Z = (z_1, z_2, z_3, z_4), Z \in \mathbb{R}^{677\times4}$$

The model was trained to minimize reconstruction error using Mean Squared Error (MSE):

$$L = \frac{1}{n}\sum(X_{scaled} - \hat{X})^2 \quad (13)$$

with Adam optimizer, 60 epochs, batch size 32, and validation split of 0.2. EarlyStopping (patience = 10) prevented overfitting, while ModelCheckpoint preserved optimal weights. The learning process followed:

$$X_{scaled} \rightarrow Autoencoder \rightarrow \hat{X}$$

This latent compression reduces the redundancy of the correlated data indicators, captures the nonlinear sustainability interactions, reduces the dimensionality from 5D to 4D prior to quantum mapping, improves noise robustness, and generates a well-structured feature space for subsequent quantum angle encoding and QSVM classification.

3.2.6 Quantum Feature Scaling, Kernel Construction, and QSVM Classification

After obtaining the latent sustainability matrix $Z \in \mathbb{R}^{677\times4}$from the autoencoder, each latent variable $z_1, z_2, z_3, z_4$was rescaled to the angular interval $[0, \pi]$, since quantum circuits require rotation parameters in bounded angular form. The transformation was performed using:

$$Z_{quantum} = \frac{Z - Z_{min}}{Z_{max} - Z_{min}} \times \pi \quad (14)$$

where $Z_{min}$and Z_{max}represent the minimum and maximum values of each latent dimension computed across the 677 samples. After scaling, all features satisfied:

$$0 \leq z_{quantum} \leq \pi$$

resulting in the quantum-ready matrix:

$$X_{quantum} \in \mathbb{R}^{677\times4}$$

This angular normalization is significant because quantum feature maps encode classical input as rotation angles. Without such restrictions on $[0, \pi]$, the stability of state preparation and kernel similarity would suffer.

The quantum embedding procedure made use of a ZZ feature map, which turns classical vectors into quantum states:

$$| \psi(x)\rangle = U_{ZZ}(x) \ | 0\rangle \quad (15)$$

where $x = (z_1, z_2, z_3, z_4)$. The similarity between two villages $x_i$and $x_j$was computed using the fidelity-based quantum kernel:

$$K(x_i, x_j) = |\langle \psi(x_i) | \psi(x_j) \rangle|^2 \quad (16)$$

This kernel quantifies the overlap of quantum states in Hilbert space, revealing nonlinear sustainability interactions that go beyond conventional linear separability.

Lastly, the Support Vector Machine classifier was fed the precomputed quantum kernel matrix. Risk labels were derived using the Environmental Sustainability Index (ESI) using quantile criteria.

$$Risk = \begin{cases} 0, & ESI \geq Q_{66\%} \\ 1, & Q_{33\%} \leq ESI < Q_{66\%} \\ 2, & ESI < Q_{33\%} \end{cases} \quad (17)$$

Thus, sustainability assessment followed the structured pipeline:

Latent Features → Angular Scaling → Quantum Embedding → Fidelity Kernel → QSVM Classification

Since it enabled the merging of nonlinear sustainability patterns identified by neural compression with a high-dimensional quantum space for better classification of village-level degradation risk, this three-stage quantum processing was vital to the proposed model.

3.2.7 Quantum circuit design 4 qubit ZZFeatureMap for Environmental index classification

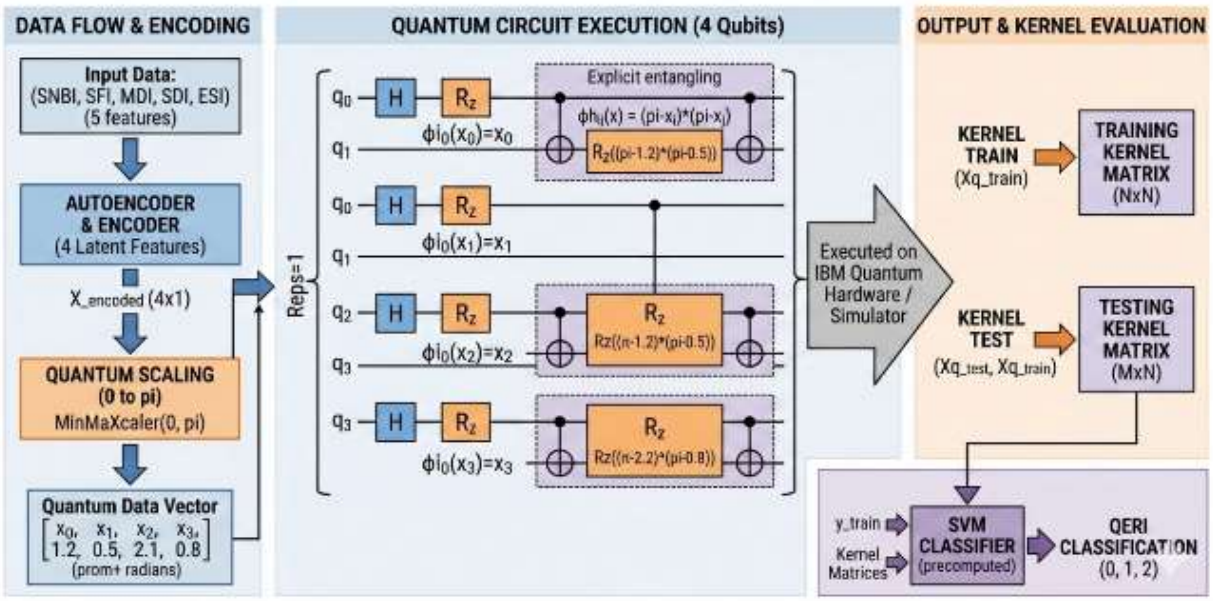

Figure 2: End-to-End Hybrid Quantum-Classical Pipeline for Environmental Risk Classification

The proposed architecture uses a 4-qubit ZZFeatureMap for encoding latent environmental information and a Deep Autoencoder for reducing the dimensionality. The model extends the accuracy of the QSVM for precise QERI classification by mapping nonlinear correlations into a high-dimensional Hilbert space using Quantum Kernel Estimation (QKE) with fidelity-based Compute Uncompute.

4. Data Visualization and Exploratory Environmental Analysis

To gain a better understanding of nutrient distribution and regional environmental variability, a variety of visualization approaches were employed.

4.1 Dataset Structural Overview

The next multi-panel figure 3 illustrates the dimensionality of the dataset (number of rows and columns), the distribution of the spatial samples in the blocks, the scientific classification of nutrition into macro, micro, and physicochemical, and the geo-administrative mapping of the villages and blocks. Figure 3 examines the data for integrity, balance, taxonomy, and traceability.

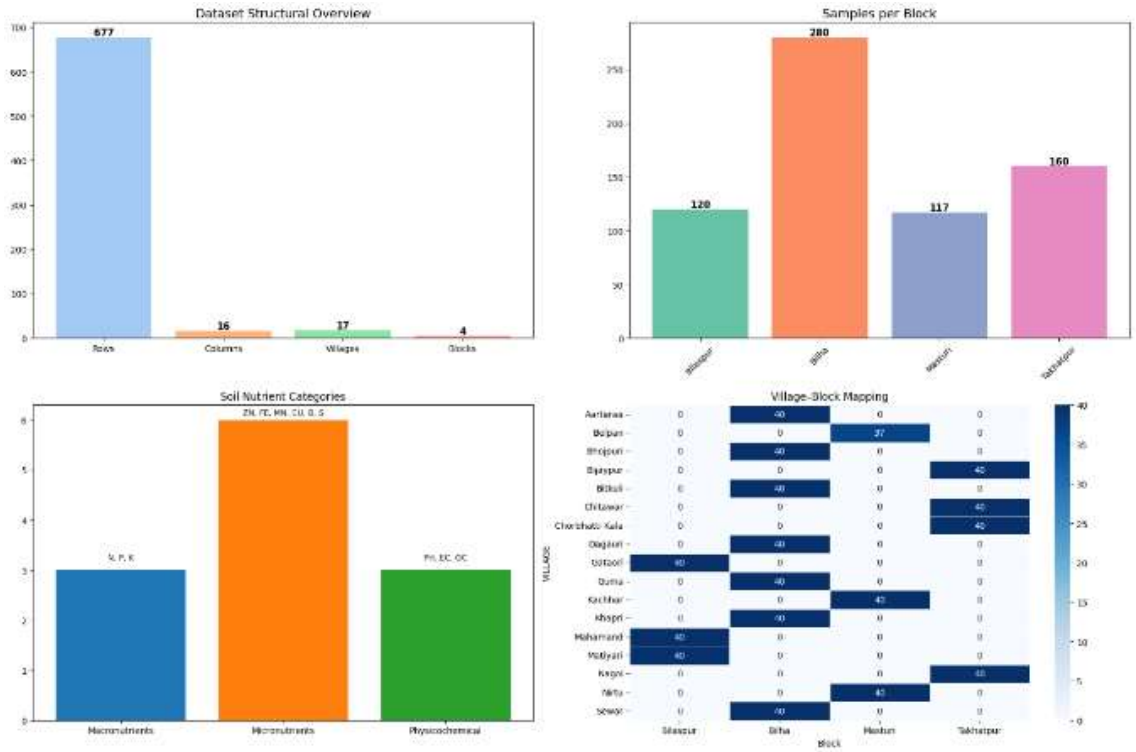

Figure 3: Comprehensive Dataset Structure and Administrative Distribution Overview

4.2 Nutrient Distribution Analysis

The nutrient distribution analysis based on the boxplot helps to understand the macro, micro, and physicochemical parameters' tendency, dispersion, skewness, and variability. It makes it easy to detect statistical dispersion, outliers, and balance in the data, ensuring that the data is stable and adequate for normalization.

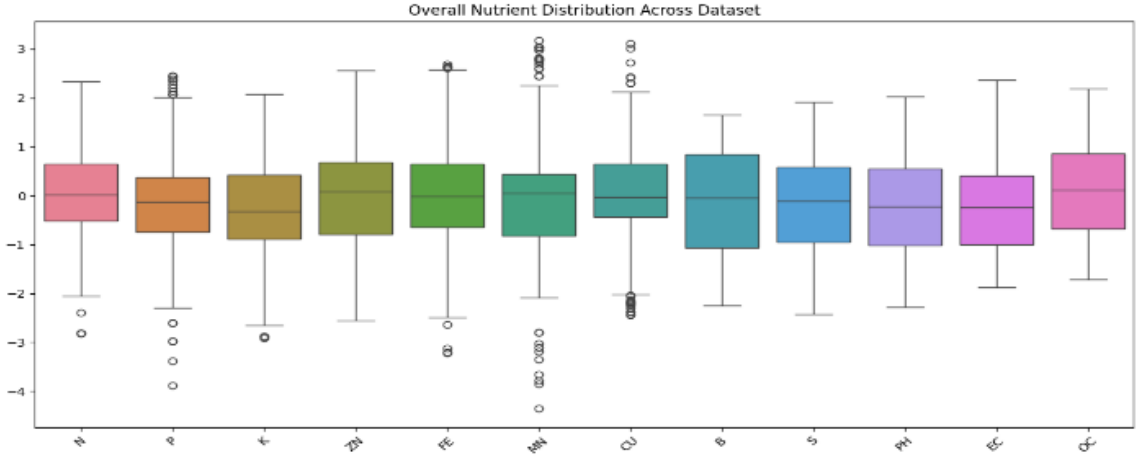

Figure 4: Nutrient Distribution Analysis

4.3 Village-Level Nutrient Benchmarking and ICAR Deviation Analysis

This dual heatmap depicts the mean nutritional levels in each hamlet, as well as their normalized deviation from ICAR norms. Deviation was computed as:

$$\text{Deviation} = \frac{\text{Observed} - \text{ICAR}_{ref}}{\text{ICAR}_{ref}} \quad (18)$$

Positive numbers imply excessive concentration, whilst negative values suggest inadequate amounts. The first panel displays aggregated nutrient intensity, while the second panel displays proportionate deviation, allowing for environmental benchmarking and community priority for sustainability.

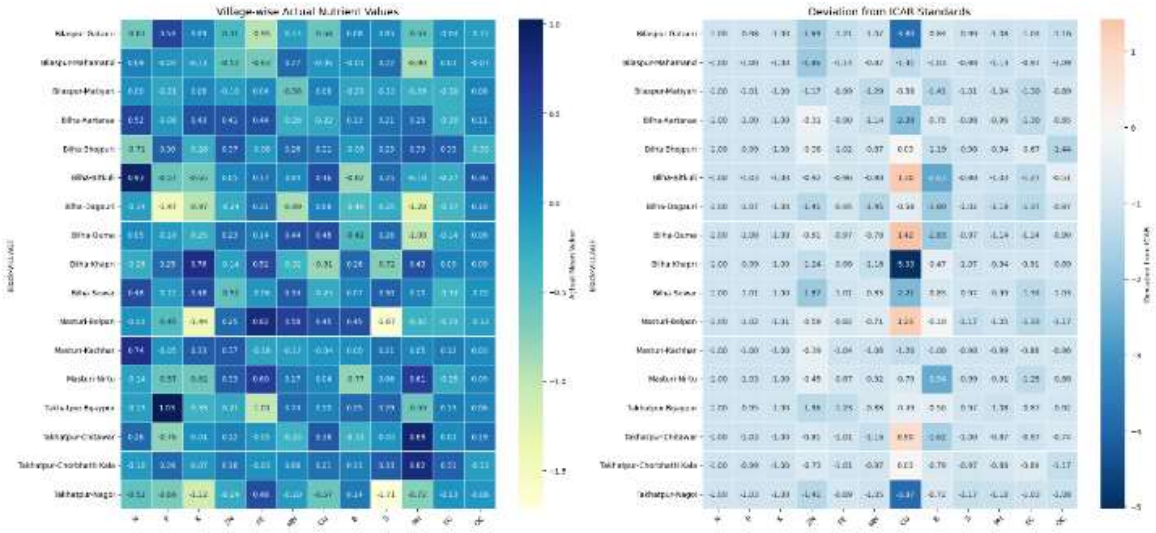

Figure 5: Shows the mean nutrient values by village and the corresponding deviation from ICAR standards across Bilaspur district.

4.4 Block-wise Nutrient Composition Analysis

A stacked bar graph was used to compare the aggregated mean nutritional content of administrative units. The total nutritional profile of each block was estimated as follows:

$$Composition_{block} = \sum_{i=1}^{n} \bar{Nutrient}_i \quad (19)$$

The code in the where statement gives the block-wise average of all macro, micro, and physicochemical parameters. The layered structure gives a more accurate assessment of the contribution of $\bar{Nutrient}_i$ and environmental variability in the Bilaspur region.

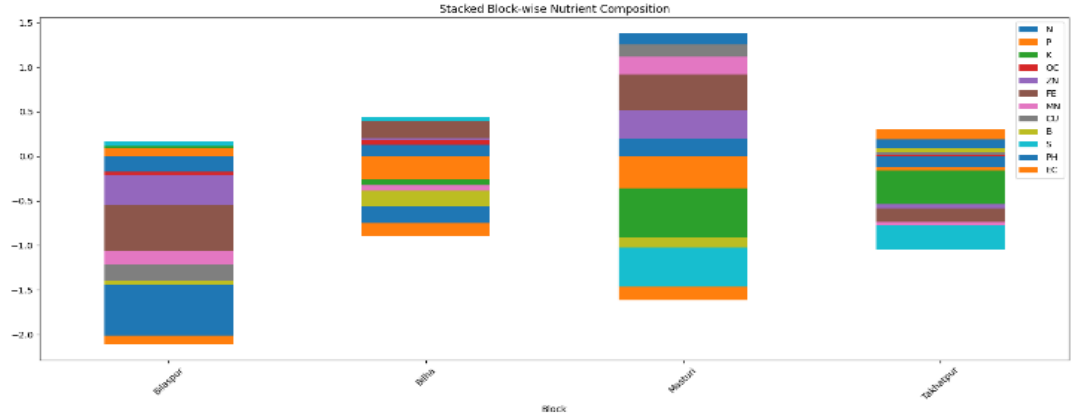

Figure 6: Stacked Bar Representation of Block-wise Mean Nutrient Composition Across Bilaspur District.

4.5 Nutrient Correlation Analysis

Pearson's correlation coefficient was calculated to measure linear interrelationships between macro, micro, and physicochemical characteristics

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sigma_x \sigma_y} \quad (20)$$

In the heatmap visualization, pairwise correlation coefficients can be seen, which indicate nutritional interdependence, multicollinearity, and environmental interaction structures that are necessary for statistical validation and dimensionality reduction.

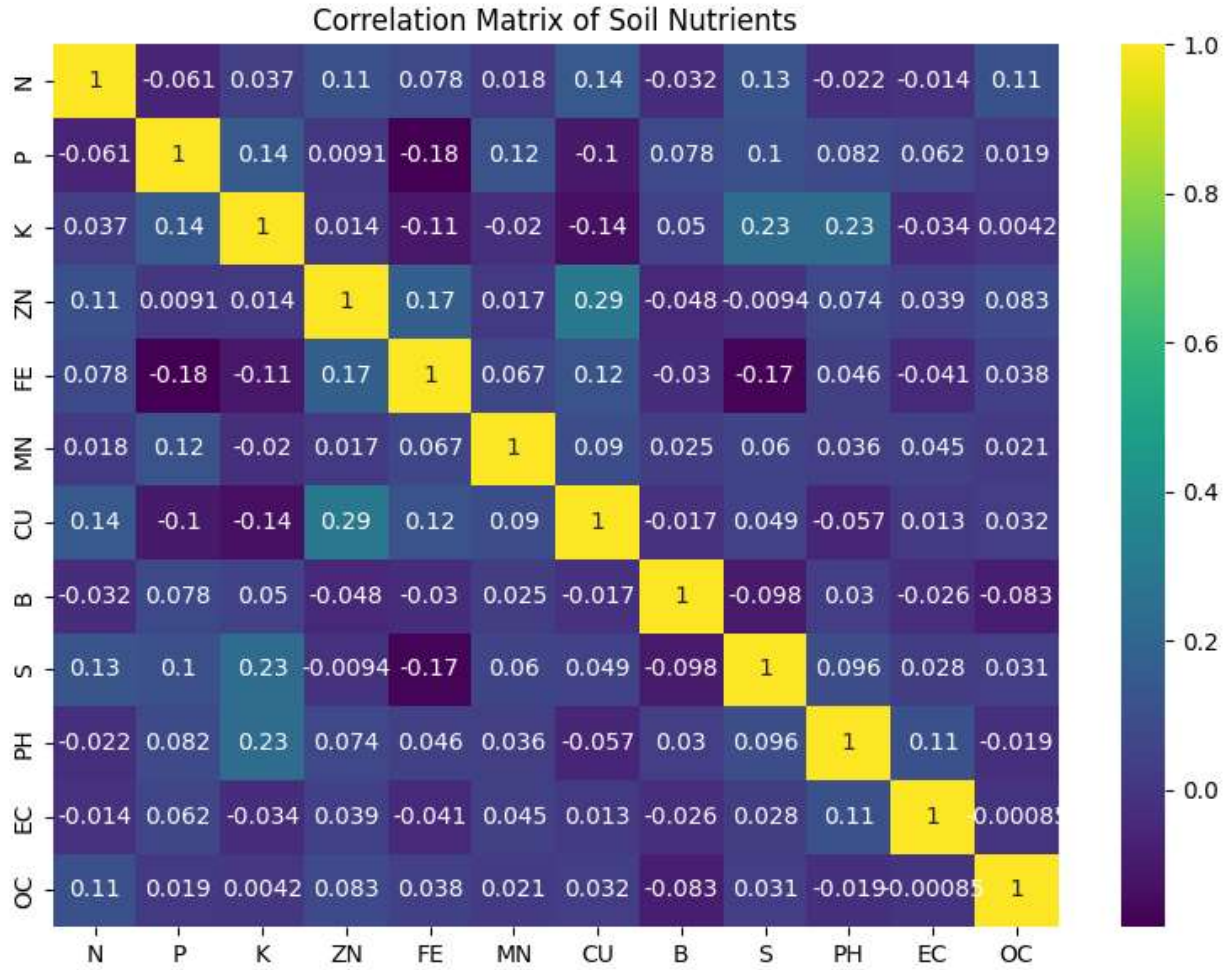

Figure 7: Pearson Correlation Heatmap Showing Interrelationships Among Soil Nutrient Parameters.

4.6 Hierarchical Cluster Analysis of Villages

Village-wise mean nutrient matrix $X$was standardized using:

$$Z = \frac{X-\mu}{\sigma} \quad (21)$$

where μ and σ indicate the nutrient-wise mean and standard deviation. The clustermap groups nutritionally comparable villages using hierarchical clustering and standardized values. This reveals nutrient-driven environmental commonalities, regional sustainability trends, and degradation processes specific to each location.

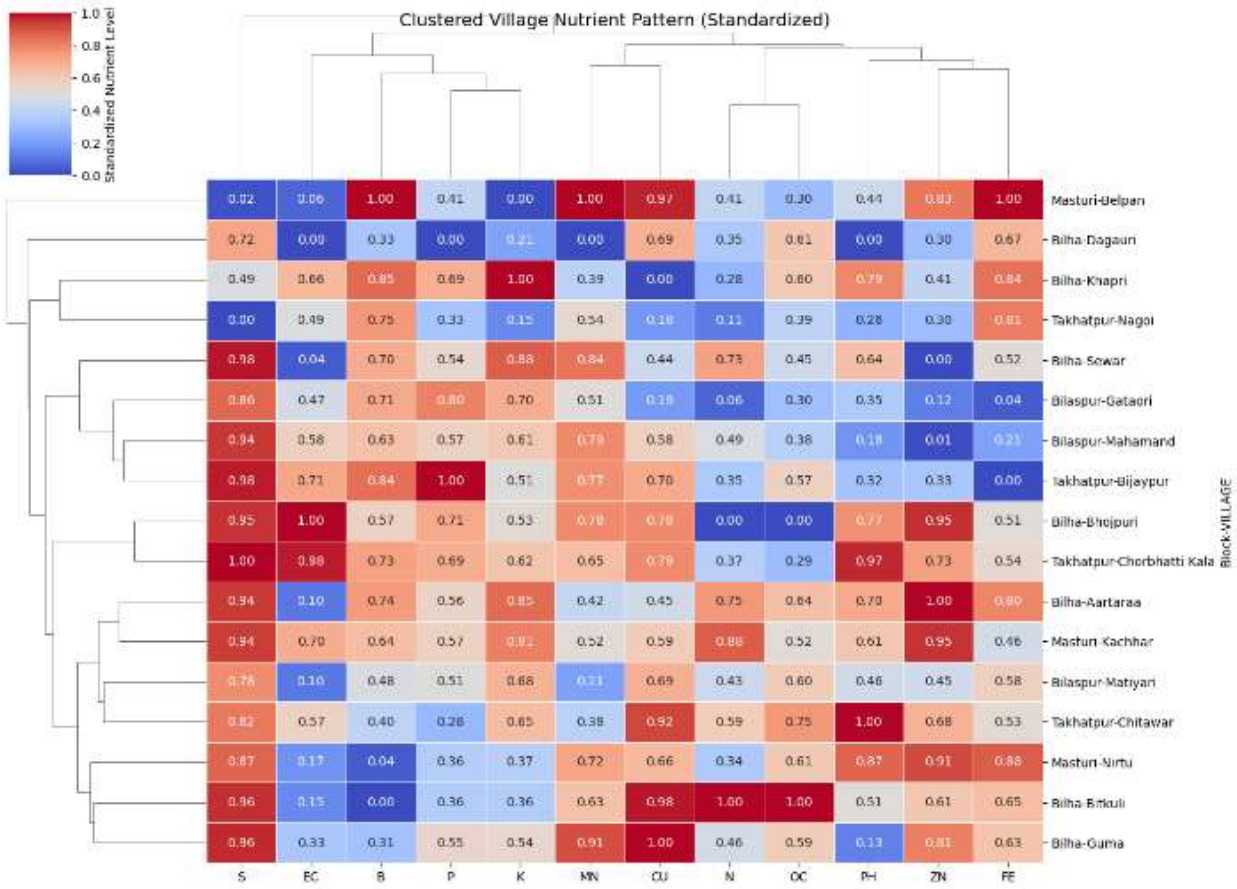

Figure 8: Hierarchical Clustering of Standardized Village Nutrient Profiles Showing Environmental Similarity Patterns Across Blocks.

5. Results and Discussion

5.1 Autoencoder Training Performance

The deep autoencoder obtained steady convergence with final training loss L=0.0013 and validation loss =0.0013. The reconstruction error was estimated using:

$$L = \frac{1}{n}\sum(X_{scaled} - \hat{X})^2 \qquad (22)$$

The small loss suggests accurate nonlinear compression of sustainability indicators into a 4-dimensional latent space without overfitting, resulting in reliable feature extraction for quantum mapping.

5.2 Risk Label Distribution

Risk categories were derived using ESI quantile thresholds:

$$Risk = \begin{cases} 0 & (High\ Sustainability) \\ 1 & (Moderate) \\ 2 & (Degraded) \end{cases}$$

Distribution:

Sustainable = 230, Moderate = 223, Degraded = 224.

The virtually balanced class distribution enhances classification robustness and reduces bias during QSVM training.

5.3 QSVM Classification Performance

The Fidelity-based QSVM achieved:

$$Accuracy = 88.23\%$$

Precision, recall, and F1-scores varied from 0.84-0.91 among classes (Eqn. 16).

Nonlinear sustainability structures were successfully recorded, indicating a strong predictive capability for assessing environmental risk at the village scale.

5.4 QML Risk Distribution

This horizontal bar plot shows the mean QERI scores per village:

$$QERI = QSVM(X_{quantum}) \quad (23)$$

It illustrates the spatial variability of vulnerability to degradation for 17 villages. A higher QERI value indicates a higher vulnerability to sustainability, which enables micro-level environmental prioritization.

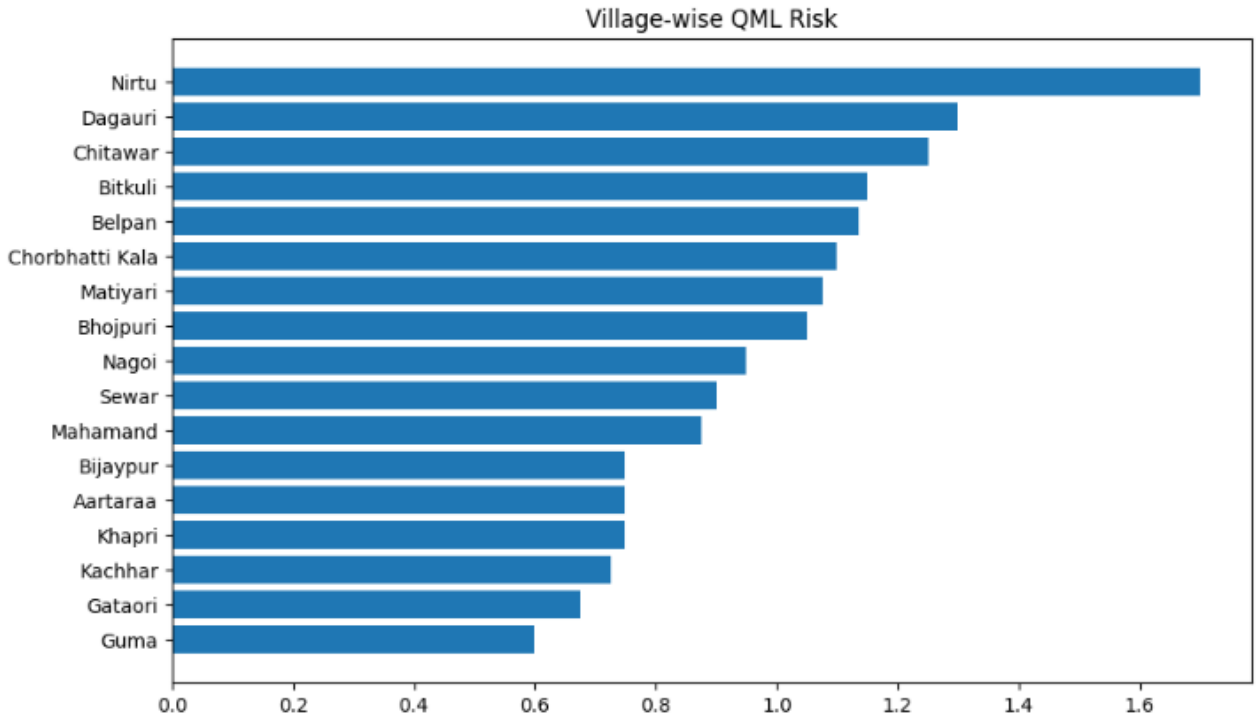

Figure 9: Village-wise QML Risk Distribution

5.5 Environmental Sustainability Distribution

The pie chart shows the relative sustainability classes within the district. The sustainability indicator displays balanced levels of sustainability, around 33% at each location, representing different environmental conditions. This underlines the need for intervention at the block level over the entire district.

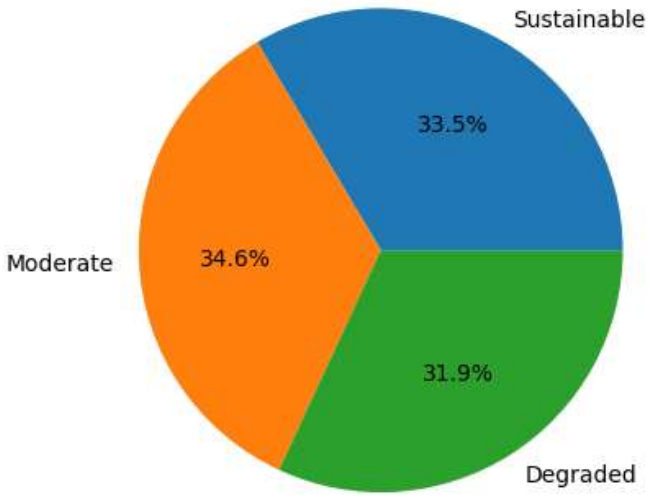

Figure 10: Regional Sustainability Distribution

5.6 Soil Environmental Sustainability

The radar chart represents mean indicator values:

$$S\bar{N}BI, S\bar{F}I, M\bar{D}I, S\bar{D}I, E\bar{S}I$$

It displays the relative magnitudes of nutrient balance, fertility index, micronutrient deficiency, soil degradation, and composite sustainability, providing a complete environmental profile in a single visual framework.

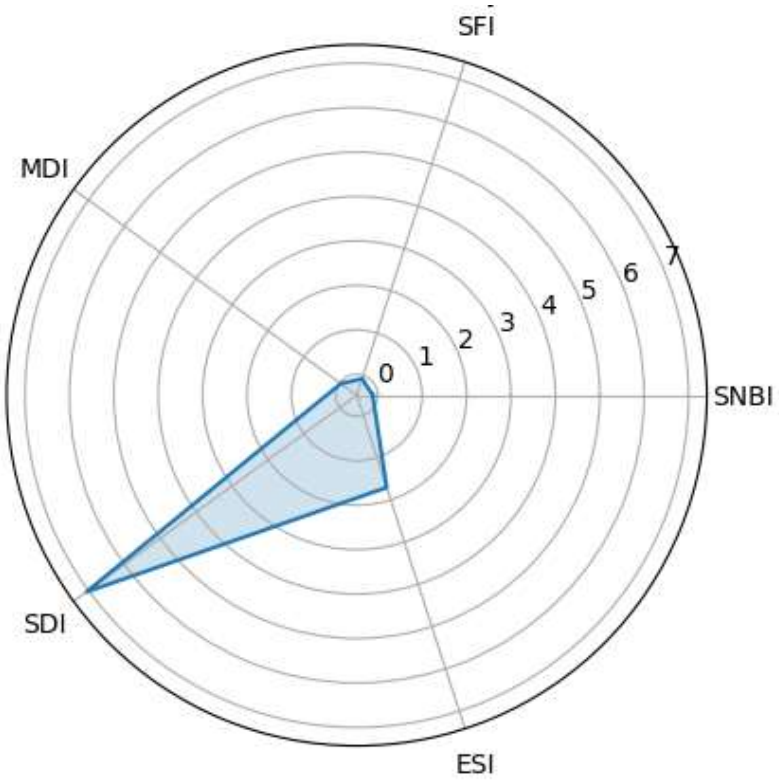

Figure 11: Soil Environmental Sustainability Radar

5.7 Environmental Parameter Interaction Matrix

The heatmap displays Pearson correlation coefficients:

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sigma_x \sigma_y} \quad (24)$$

It reveals interdependence among sustainability indicators, hence validating the autoencoder's nonlinear interactions. Prior to quantum kernel embedding, strong correlations facilitate dimensional compression.

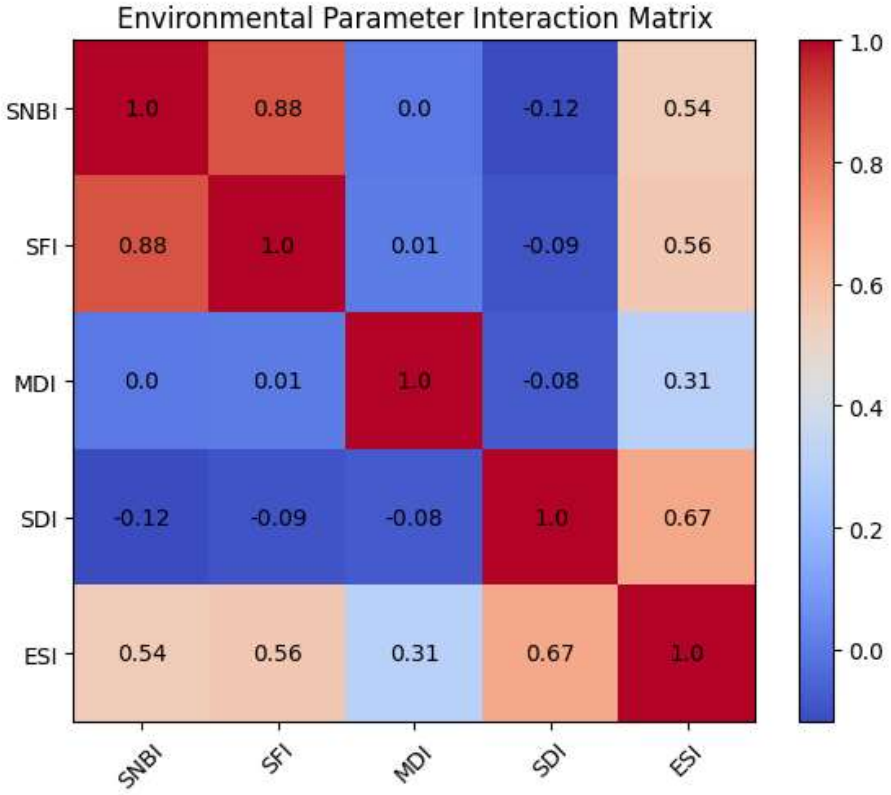

Figure 12: Environmental Parameter Interaction Matrix

6. Conclusion

This study proposes a quantum-classical hybrid analytical model for assessing soil sustainability and degradation risk in villages in Bilaspur. The model utilizes deep autoencoder networks for feature learning in a compact form from environmental parameters and a 4-qubit ZZFeatureMap for projecting features in a high-dimensional Hilbert space. Quantum feature encoding allows the model to incorporate complex and non-linear interactions among soil nutrients and physicochemical parameters, which may not be considered by traditional machine learning algorithms.

Quantum Support Vector Machine (QSVM) achieved an accuracy of about 88.23%. This shows that QSVM can be used to distinguish between sustainability conditions in different locations. The study shows significant differences in soil health in different villages, which highlights the drawbacks of using a single soil management practice for different villages. This quantum-classical model can thus pave the way for using localized and data-driven decision-making in agriculture. The study thus proposes a quantum-classical model for precision agriculture systems that can incorporate parameters such as climate change, crop productivity, and dynamic land management strategies in the future.

7. Acknowledgement

The authors would like to express their sincere gratitude to Dr. Amit Kumar Chandanan for his valuable guidance, support, and supervision throughout the research work.

The authors also acknowledge all individuals and institutions who directly or indirectly contributed to the completion of this study.

8. References

1. R. Lal, "Soil health and climate change," Science, vol. 304, no. 5677, pp. 1623–1627, 2004, doi: 10.1126/science.aag0809.
2. J. W. Doran and M. R. Zeiss, "Soil health and sustainability: Managing the biotic component of soil quality," Applied Soil Ecology, vol. 15, no. 1, pp. 3–11, Aug. 2000, doi: 10.1016/S0929-1393(00)00067-6.
3. N. K. Jat et al., "Comparative fertilization effects on maize productivity under conservation and conventional tillage on sandy soils in a smallholder cropping system in Zimbabwe," Field Crops Research, vol. 218, pp. 106–114, Apr. 2018, doi: 10.1016/j.fcr.2018.01.014.
4. E. Y. Obsie, H. Qu, and F. Drummond, "Wild blueberry yield prediction using a combination of computer simulation and machine learning algorithms," Computers and Electronics in Agriculture, vol. 178, p. 105778, Nov. 2020, doi: 10.1016/j.compag.2020.105778.
5. C. M. G. Reis, J. Raimundo, and M. M. Ribeiro, "Assessment of genetic diversity in Opuntia spp. Portuguese populations using SSR molecular markers," Agronomy, vol. 8, no. 4, p. 55, 2018, doi: 10.3390/agronomy8040055.
6. M. Wolff, "Taking one step further – Advancing the measurement of green and blue area accessibility using spatial network analysis," Ecological Indicators, vol. 126, p. 107665, Jul. 2021, doi: 10.1016/j.ecolind.2021.107665.
7. A. Guo, W. Huang, Y. Dong, H. Ye, H. Ma, B. Liu, W. Wu, Y. Ren, C. Ruan, and Y. Geng, "Wheat yellow rust detection using UAV-based hyperspectral technology," Remote Sensing, vol. 13, no. 1, p. 123, Jan. 2021, doi: 10.3390/rs13010123.
8. M. Schuld and F. Petruccione, Quantum Machine Learning: An Introduction. Cham, Switzerland: Springer, 2018, doi: 10.1007/978-3-319-96424-9.
9. V. Havlíček et al., "Supervised learning with quantum-enhanced feature spaces," Nature, vol. 567, pp. 209–212, Mar. 2019, doi: 10.1038/s41586-019-0980-2.
10. M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models," npj Quantum Information, vol. 5, p. 45, 2019, doi: 10.1038/s41534-019-0138-4.
11. E. Farhi and H. Neven, "Classification with quantum neural networks on near term processors," arXiv:1802.06002, 2018. [Online]. Available: https://arxiv.org/abs/1802.06002.
12. Indian Council of Agricultural Research (ICAR), Soil Health Management Manual. New Delhi, India: Govt. of India, 2017. [Online]. Available: https://icar.org.in.
13. B. S. Das et al., "Soil health and its relationship with food security and human health to meet the sustainable development goals in India," Soil Security, vol. 8, p. 100071, Sep. 2022, doi: 10.1016/j.soisec.2022.100071.
14. X. Yang, N. Bao, W. Li, S. Liu, Y. Fu, and Y. Mao, "Soil nutrient estimation and mapping in farmland based on UAV imaging spectrometry," Sensors, vol. 21, no. 11, p. 3919, Jun. 2021, doi: 10.3390/s21113919.

15. X. Vasques, H. Paik, and L. Cif, "Application of quantum machine learning using quantum kernel algorithms on multiclass neuron M-type classification," Scientific Reports, vol. 13, p. 11541, Jul. 2023, doi: 10.1038/s41598-023-38558-z.
16. [16] J. Schnabel and M. Roth, "Quantum kernel methods under scrutiny: A benchmarking study," Quantum Machine Intelligence, vol. 7, p. 58, Apr. 2025, doi: 10.1007/s42484-025-00273-5.