Agrociencia

SOIL QUALITY FORECASTING WITH OPTIMAL FEATURE SELECTION AND EXTENDED CROSS-STAGE PYRAMID NETWORK

Yamuna **Vilvanathan**[1], Jeevaa **Katiravan**[2], Visu **Pandurangan**[3*], Smitha Sarojam **Ponnaiyan**[4]

[1]Sri Sairam College of Engineering. Department of Computer Science and Engineering. Anekal, Bangalore, Karnataka 562106, India.
[2]Velammal Engineering College. Department of Information Technology. Chennai, Tamil Nadu 600066, India.
[3]Velammal Engineering College. Department of Artificial Intelligence and Data Science. Chennai, Tamil Nadu 600066, India.
[4]Velammal Engineering College. Department of Computer Science and Engineering. Chennai, Tamil Nadu 600066, India.
* Author for correspondence: dr.visu.p@gmail.com

ABSTRACT

Soil quality is essential for sustainable agriculture. Nonetheless, inadequate irrigation methods, improper fertilizer use, and over-cultivation reduce soil quality, thereby decreasing soil fertility. Precise soil quality prediction is crucial for improving agricultural practices. Conventional deep learning models often encounter problems related to superfluous features, high computational demands, and inaccurate predictions. In this work, an innovative deep learning architecture integrating optimal feature selection was proposed to address these challenges. Initially, the Adaptive Parrot Optimization (AdPo) method was used to identify the most relevant features from pre-processed soil data. The Extended Cross Stage Pyramid Network (ExCSP_Net) was introduced to improve soil quality prediction. This network integrates a gated recurrent unit (GRU)-based attention module into the main pathway of the Cross Stage Partial (CSP) model to capture long-range dependencies and emphasize relevant information. In addition, a stacked autoencoder was incorporated before the feature-sharing stage in the short path of the CSP model to reduce dimensionality and generate meaningful representations. The AdPo+ExCSP_Net model demonstrated outstanding performance, achieving an accuracy of 98.58 %, recall of 98.09 %, precision of 98.32 %, F1-score of 98.15 %, Mean Absolute Error (MAE) of 0.53, Root Mean Square Error (RMSE) of 0.65, and coefficient of determination (R^2) of 0.99. These findings highlight the effectiveness of the proposed methodology for accurate soil quality prediction and the promotion of sustainable agricultural practices.

Keywords: adaptive weighting, missing data imputation, cross stage pyramid network, optimal feature selection, soil quality.

Citation: Vilvanathan Y, Katiravan J, Pandurangan V, Ponnaiyan SS. 2026. Soil quality forecasting with optimal feature selection and extended cross-stage pyramid network. **Agrociencia.** https://doi.org/ 10.47163/agrociencia.v60i4.3465

Editor in Chief:
Dr. Fernando C. Gómez Merino

Received: September 13, 2025.
Approved: May 15, 2026.
Published in Agrociencia: June 01, 2026.

INTRODUCTION

Agricultural management is essential for ensuring the productivity and sustainability of farming systems. It involves the control, monitoring, and planning of agricultural activities, including resource allocation, irrigation, pest management, livestock care, and crop production (Li *et al*., 2025). Effective agricultural management is necessary to meet the growing demand for food, address environmental transformation, and protect natural resources (Song *et al*., 2024). Agricultural yields can be enhanced by optimizing farming practices while minimizing costs and environmental impacts (Safaie *et al*., 2023).

One of the most important aspects of agricultural management is soil quality management. Soil is a fundamental component of productive agriculture because it provides essential nutrients that support crop production (Barathkumar *et al*., 2025). Soil with higher organic matter content and an appropriate nutrient composition promotes the development of healthier plants, thereby improving crop productivity. In contrast, poor-quality soil limits crop productivity due to nutrient deficiencies and increases crop susceptibility to pests and diseases (Huang *et al*., 2023).

Soil quality is also critical for the long-term sustainability of farming. Poor irrigation management, excessive fertilizer use, and over-farming negatively affect soil quality through fertility loss, erosion, and soil degradation (Raza *et al*., 2023). Practices such as reduced tillage, cover cropping, organic farming, and crop rotation contribute to improved agricultural management by preserving and enhancing soil health (Omondiagbe *et al*., 2023). Effective agricultural management helps maintain soil fertility, enhance production, reduce the need for chemical inputs, and support sustainable agricultural productivity (Ghani *et al*., 2024).

Traditional soil evaluation methods involve the manual collection of soil samples from various locations, followed by chemical and physical analyses in laboratories to assess soil health based on the obtained results (Chaudhry *et al*., 2024). Accordingly, traditional soil quality forecasting relies on field observations, laboratory testing, physical soil sampling, historical data analysis, geographic surveys, and soil mapping to evaluate soil characteristics such as moisture levels, nutrient content, and pH levels, as well as to predict soil quality across larger areas (Suwardi *et al*., 2023; Pant *et al*., 2024). However, these methods face major challenges because they are expensive, labor-intensive, and time-consuming, particularly when large geographical areas are considered (Parewai and Köppen, 2024). In addition, soil properties can vary considerably over short distances, making it difficult to accurately capture soil characteristics across different regions (El Behairy *et al*., 2024). Spatial variability and the lack of real-time data further reduce the efficiency of traditional soil quality forecasting methods for continuous soil monitoring and agricultural management (Divya *et al*., 2024).

Deep learning-based methods for predicting soil quality provide an effective solution to many of the challenges associated with conventional techniques by using advanced algorithms and large datasets to achieve greater speed and accuracy (Chen *et al*., 2023). Using historical records, remote sensing data, and satellite imagery, deep learning approaches can analyze complex soil characteristics without requiring extensive manual sampling (Selvanarayanan *et al*., 2024). In addition, these models are highly scalable, enabling more efficient soil quality assessment across large geographical regions (Shahzad *et al*., 2024). However, certain limitations remain. The accuracy of deep learning models depends heavily on the availability and quality of the training data (Shahare *et al*., 2024). Poor-quality or limited datasets may result in unreliable predictions, thereby increasing the complexity of agricultural management (Kalyani and Kolla, 2024). To address these limitations, a deep learning-based framework for forecasting soil quality with optimal feature selection was proposed in this study (Sumathi *et al*., 2023).

The primary objective of this research was to develop an efficient framework for forecasting soil quality that can improve prediction accuracy and convergence performance. For this purpose, the Adaptive Parrot Optimization (AdPo) algorithm was designed for feature selection to enhance the convergence rate in identifying the global optimal solution. In this approach, adaptive weighting was incorporated into the exploration stage of the parrot optimization algorithm to improve convergence efficiency during the search process. Furthermore, soil quality forecasting was performed using the proposed Extended Cross Stage Pyramid Network (ExCSP_Net) to improve forecasting accuracy. In the ExCSP_Net architecture, a gated recurrent unit (GRU)-based attention module was integrated into the main path of the Cross Stage Partial (CSP) model to capture important feature dependencies, while a stacked autoencoder was incorporated before the feature-sharing stage in the short path of the CSP model to reduce feature dimensionality and generate meaningful representations.

MATERIALS AND METHODS

Soil quality forecasting using optimal feature selection and deep learning-based prediction was introduced in this study (Figure 1). Initially, the soil data were acquired

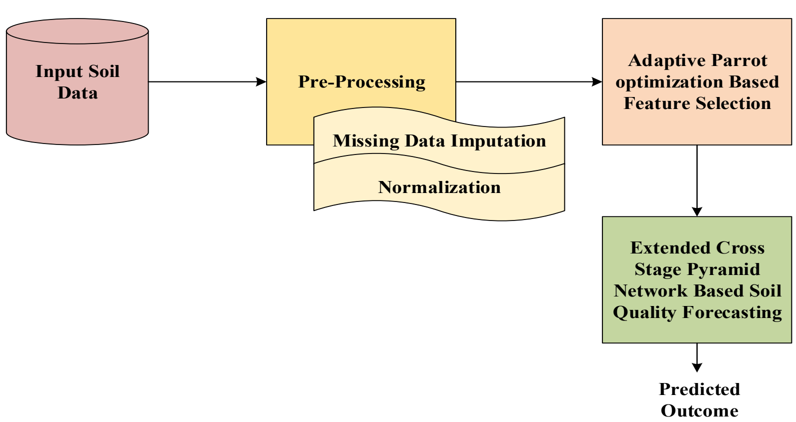

Figure 1. General workflow of the proposed soil quality forecasting framework.

from a publicly available dataset and pre-processed using missing data imputation and normalization techniques. Subsequently, significant features were selected from the preprocessed data using the Adaptive Parrot Optimization (AdPo) algorithm. Finally, soil quality prediction was performed using the Extended Cross Stage Pyramid Network (ExCSP_Net). For the development of ExCSP_Net, a gated recurrent unit (GRU)-based attention module was incorporated into the main path of the Cross Stage Partial (CSP) model, and a stacked autoencoder was integrated before the feature-sharing stage in the short path of the CSP model to minimize feature dimensionality.

Data acquisition

The input data for the proposed soil quality forecasting framework were obtained from the publicly available Soil Fertility Dataset available on Kaggle (https://www.kaggle.com/datasets/rahuljaiswalonkaggle/soil-fertility-dataset). The dataset contained 38 600 soil samples with multiple numerical attributes, including pH, nitrogen (N), phosphorus (P), potassium (K), organic carbon, moisture content, electrical conductivity, and other fertility-related indicators, resulting in a total of N predictor variables (Oukhattar *et al.*, 2025). The response variable corresponded to soil fertility status, representing the soil quality class indicating soil suitability for crop growth (Ziyadullaev *et al.*, 2024).

The proposed Adaptive Parrot Optimization with Extended Cross Stage Pyramid Network (AdPo+ExCSP_Net) model was implemented in the Python programming language on a Windows 10 operating system with 8 GB RAM. Conventional approaches, including Artificial Neural Network (ANN), Extreme Gradient Boosting (XGBoost), Random Forest (RF), Improved Soil Quality Prediction-Deep Learning (ISQP-DL), and the proposed model without AdPo, were used for comparative analysis. Model performance was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), coefficient of determination (R^2), processing time, F-score, recall, precision, and accuracy.

Pre-processing

Pre-processing based on missing data imputation and normalization was applied in the proposed soil quality forecasting model. Missing data imputation was performed to enable the effective use of incomplete datasets by estimating missing values while preserving the overall structure and richness of the data, preventing information loss and improving prediction accuracy. Missing values were estimated using the K-Nearest Neighbors (KNN)-based imputation method, in which predictions were generated from the values of the nearest neighbors in the dataset. The KNN-based approach considered the relationships among features and provided accurate imputation for soil data. The steps involved in KNN-based missing data imputation were as follows.

Step 1: The Euclidean distance between data points was calculated as:

$$D(X_u, X_v) = \sqrt{\sum_{s=1}^{S} \left(X_{u,s} - X_{v,s}\right)^2}$$

where the data points are represented as $X_{u,s}$ and $X_{v,s}$, respectively, the total number of data points is represented as S, and the distance between them is denoted as $D(X_u, X_v)$. The S nearest neighbors were selected based on the calculated distance.

Step 2: Missing data imputation was performed by estimating the mean of the selected S data points. Following missing data imputation, normalization was performed to transform the data into a standard form.

Normalization

Z-score-based normalization was applied to standardize the data and was expressed as:

$$I_z = \frac{x - Avg}{SD}$$

where I_z represents the normalized output, Avg is the average value, SD is the standard deviation, and x is the input value. Significant feature selection was subsequently performed from the preprocessed data using the proposed AdPo algorithm.

Feature selection

AdPo-based feature selection was used to identify the optimal feature subset from the pre-processed data. The proposed AdPo algorithm was designed to improve the convergence rate of the optimization process in identifying the global optimal solution. In this approach, adaptive weighting was incorporated into the exploration stage of the parrot optimization algorithm to enhance convergence efficiency during the search process. Forecasting accuracy was considered a fitness function for selecting the optimal feature subset. The population size, represented by L, denotes the number of parrots (solutions) in the swarm, whereas $Ymax$ is the maximum number of iterations used to update the positions of the solutions.

The initial position of the parrot was calculated as:

$$P_r^0 = D + u(0,1) \cdot (X - D)$$

where D and X are the lower and upper boundaries of the search space, respectively. The term $u(0,1)$ is a continuous uniform random variable within the interval (0,1), which introduces randomness during initialization, and P_r^0 represents the initial position of the *r*th parrot.

The position updating process was expressed as:

$$P_r^{Y+1} = (P_r^Y - P_{best}) \cdot Q(n) + u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right)^{2 \cdot \frac{y}{Y_{max}}} \cdot P_{Avg}^Y$$

where P_r^Y represents the current position of the *r*th parrot, $P_r^{(Y+1)}$ is the updated position after iteration, and P_{best} indicates the best position identified by the swarm, corresponding to the optimal solution obtained thus far. The term $Q(n)$ represents the Lévy flight distribution, which enables long exploratory movements during the search process and facilitates the identification of new solution regions. The variable $u(0,1)$ is a continuous uniform random variable within the interval (0,1), y represents the current iteration, and Y_{max} indicates the maximum number of iterations. In addition, P_{Avg}^Y is the average position of the swarm population, which guided the solutions according to the general movement trend of the swarm.

The solution updating mechanism of the proposed AdPo algorithm consisted of two major stages, namely exploration and exploitation, represented by $(P_r^Y - P_{best}) \cdot Q(n)$ and $u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right)^{2 \cdot \frac{y}{Y_{max}}} \cdot P_{Avg}^Y$, respectively.

The average position of the swarm at iteration Y was calculated as:

$$P_{Avg}^Y = \frac{1}{L} \sum_{d=1}^{L} P_{dY}$$

where L is the total population size and P_{dY} is the position of the *d*th parrot at iteration Y.

Lévy flight

This approach was used as a random walk mechanism in which most movements were small, while occasional large jumps allowed extensive exploration of the search space. The Lévy flight distribution was defined as:

$$Q(n) = \frac{\delta \cdot \varepsilon}{|s|}$$

where $\delta \sim N(0,n)$ and $s \sim N(0,n)$ represent normally distributed random variables, and ε was calculated as:

$$\varepsilon = \left(\frac{\Gamma(1+\lambda) \cdot sin\left(\frac{\pi \cdot \lambda}{2}\right)}{\Gamma\left(\frac{1+\lambda}{2}\right) \cdot \lambda \cdot 2^{\frac{1}{\lambda}}} \right)^{\frac{1}{\lambda}}$$

where λ represents the parameter controlling the shape of the distribution and was set to 1.5. Lévy flights were effective for both local and global search processes, thereby enabling the optimization algorithm to explore both nearby and distant regions of the solution space.

To improve the foraging behavior of the parrots, an adaptive weighting strategy was incorporated and expressed as:

$$Ad_w = P_r^{Y+1} + F\left((P_r^Y)_{good} - (P_r^Y)\right)$$

where Ad_w represents the adaptive weighting factor, $P_r^{(Y+1)}$ denotes the solution obtained in the current iteration, and P_r^Y is the solution obtained by the parrot search agent in the previous iteration. The term $(P_r^Y)_{good}$ indicates the best solution identified in the previous iteration, while F is the controlling parameter.

After incorporating adaptive weighting, the updated solution of the proposed AdPo algorithm was formulated as:

$$(P_r^{Y+1})_{AdPo} = Ad_w * (P_r^{Y+1})_{parrot}$$

Substituting the adaptive weighting expression into the position updating process yielded:

$$(P_r^{Y+1})_{AdPo} = \left[P_r^{Y+1} + F\left((P_r^Y)_{good} - (P_r^Y)\right)\right] *$$

$$\left[(P_r^Y - P_{best}) \cdot Q(n) + u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right)^{2 \cdot \frac{y}{Y_{max}}} \cdot P_{Avg}^Y\right]$$

The modified solution enhanced the exploration capability of the optimization process while improving convergence speed for identifying the global optimal solution.

Staying behavior. This represents the tendency of parrots to temporarily remain near their owner after movement. This behavior was modeled as a sudden movement toward the owner followed by a stationary phase and was expressed as:

$$P_r^{Y+1} = P_r^Y + P_{best} \cdot Q(n) + u(0,1) \cdot ones(1,n)$$

where $P_r^{(Y+1)}$ represents the updated position of the rth parrot, P_r^Y denotes the current position of the rth parrot, and $P_{best} \cdot Q(n)$ is the movement toward the owner modeled using the Lévy distribution. The term $u(0,1) \cdot ones(1,n)$ represents the random stopping behavior near the owner.

Communicating behavior. This approach was used to model the natural social interactions of parrots, which are highly social animals. This allows individual parrots to exchange information and update their positions according to interactions within the swarm, thereby improving optimization performance. The communication process, either through movement toward the flock or through remote interaction, was modeled as:

$$P_r^{Y+1} = \begin{cases} 0.2 \cdot u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right) \cdot \left(P_r^Y - P_{Avg}\right) & if\ Q \leq 0.5 \\ 0.2 \cdot u(0,1) \cdot exp\left(\frac{y}{u(0,1) \cdot Y_{max}}\right) & if\ Q > 0.5 \end{cases}$$

where $P_r^{(Y+1)}$ represents the updated position of the rth parrot at iteration Y+1, and P_r^Y is the current position of the rth parrot at iteration Y. The term $u(0,1)$ generates a random number between 0 and 1 to introduce variability and randomness into the communication process. In addition, Q is a random value between 0 and 1 that determines whether the parrot moved toward the flock or communicated without movement.

Fear of strangers. This behavior was incorporated to simulate the natural tendency of parrots to avoid unfamiliar environments and remain close to safe regions. It ensures that the optimization process avoided unknown or less optimal regions of the search space and moved toward safer and more optimal solutions. The fear of strangers behavior was expressed as:

$$P_r^{Y+1} = P_r^Y + u(0,1) \cdot cos\left(0.5\pi \cdot \frac{y}{Y_{max}}\right) - cos(u(0,1) \cdot \pi) \cdot \left(\frac{y}{Y_{max}}\right)^2 \cdot (P_r^Y - P_{best})$$

where $u(0,1) \cdot cos\left(0.5\pi \cdot \frac{y}{Y_{max}}\right)$ describes the movement of the parrot toward a safe region. The term $cos(u(0,1) \cdot \pi) \cdot \left(\frac{y}{Y_{max}}\right)^2 \cdot (P_r^Y - P_{best})$ models the avoidance of unfamiliar or unfavorable regions in the search space. The variable $u(0,1)$ represents a random value within the interval [0,1], P_{best} is the best position identified by the swarm,

y is the current iteration, Y_{max} is the maximum number of iterations, and P_r^Y indicates the current position of the rth parrot.

Feasibility estimation

The feasibility of the proposed AdPo algorithm was evaluated based on forecasting accuracy and was calculated as:

$$Fit = \frac{TP + TN}{TP + TN + FP + FN}$$

where *Fit* represents the fitness value, *TP* denotes the true positives, *TN* the true negatives, *FP* the false positives, and *FN* the false negatives.

Termination

The algorithm terminated when the global optimal solution was achieved or when the maximum number of iterations was completed. Based on the solution obtained using the AdPo algorithm (Algorithm 1), the optimal feature subset was selected to eliminate redundant features.

Algorithm 1. Pseudo-code of the proposed Adaptive Parrot Optimization (AdPo) algorithm for optimal feature selection.

Start
Initialize the parameters
Initialize the position of parrot search agent
Estimate the fitness for parrot search agent
Sort and update the solutions
Choose the strategy randomly
{
Update the solution based on foraging behavior

$$(P_r^{Y+1})_{AdPo} = \left[P_r^{Y+1} + F\left((P_r^Y)_{good} - (P_r^Y)\right)\right] * \left[(P_r^Y - P_{best}) \cdot Q(n) + u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right)^{2 \cdot \frac{y}{Y_{max}}} \cdot P_{Avg}^Y\right]$$

Update the solution based on staying behavior

$$P_r^{Y+1} = P_r^Y + P_{best} \cdot Q(n) + u(0,1) \cdot ones(1,n)$$

Update the solution based on communicating behavior

$$P_r^{Y+1} = \begin{cases} 0.2 \cdot u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right) \cdot \left(P_r^Y - P_{Avg}\right) & if\ Q \leq 0.5 \\ 0.2 \cdot u(0,1) \cdot exp\left(\frac{y}{u(0,1) \cdot Y_{max}}\right) & if\ Q > 0.5 \end{cases}$$

Algorithm 1. Continued

Update the solution based on fear of strangers

$$P_r^{Y+1} = P_r^Y + u(0,1) \cdot cos\left(0.5\pi \cdot \frac{y}{Y_{max}}\right) - cos(u(0,1) \cdot \pi) \cdot \left(\frac{y}{Y_{max}}\right)^2 \cdot (P_r^Y - P_{best})$$

}
Check the feasibility
Check $y > Y^{max}$
$Y = y + +$
Return the best solution
end

Extended cross stage pyramid network-based soil quality forecasting

Soil quality forecasting was performed using ExCSP_Net to improve forecasting accuracy (Figure 2). For the development of ExCSP_Net, a GRU-based attention module was incorporated into the main path of the CSP model, and a stacked autoencoder was integrated before the feature-sharing stage in the short path of the CSP model to minimize feature dimensionality. The CSP network improved the learning efficiency of deep neural networks by reducing computational complexity and enhancing gradient flow. The CSP model operated by dividing feature maps into two parts across different stages and subsequently merging them within the network to improve feature representation learning. In addition, it combined pyramid structures with cross-stage feature fusion to process multiple scales in the input data for soil quality forecasting.

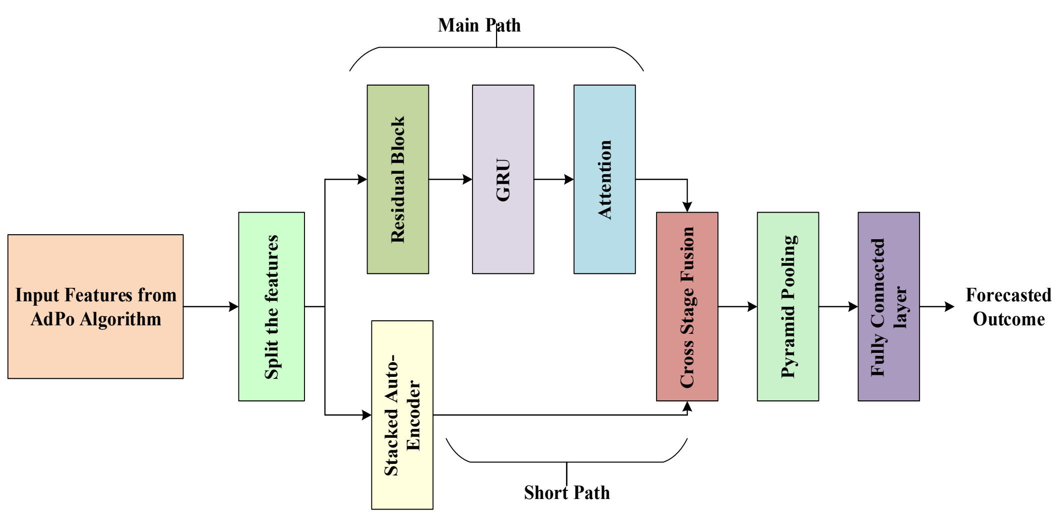

Figure 2. Architecture of the proposed Extended Cross Stage Pyramid Network (ExCSP_Net) for soil quality forecasting.

The features selected by the AdPo algorithm E were input into the ExCSP_Net model's input layer and the feature extraction module designed with both main and short paths. Initially, the features were split to facilitate their passage through these main and short paths. The splitting of features was expressed as:

$$E_M, E_S = Split(E)$$

where *EM* represents the features forwarded to the main path, and *ES* represents the features provided to the short path of the ExCSP_Net model.

The main path of the ExCSP_Net model consisted of multiple residual blocks for extracting complex features. In the proposed architecture, a GRU_Attn layer was integrated with the residual blocks to extract spatial and temporal features for improving forecasting accuracy. Initially, the features from the main path were provided to the residual block for spatial feature extraction. The residual block utilized residual connections to address the vanishing gradient problem and enabled effective training of very deep neural networks. The residual block was defined as:

$$B(e) = C(e, \{D_r\}) + e$$

where *B(e)* represents the output of the residual block, $C(e,\{D_r\})$ is the output of the convolutional layers within the block corresponding to the learned residual representation, and *e* is the input feature map provided through the skip connection. The extracted features were then sent to the activation layer, which introduced non-linearity into the learned representations. The use of residual connections allowed these features to propagate through deeper layers without losing the original spatial information. As a result, the residual blocks enabled the network to maintain fine-grained spatial structures from earlier layers while also learning higher-level feature abstractions.

The deep features extracted through the residual block were subsequently provided to the GRU module for temporal feature extraction, enabling the acquisition of long-term dependent features. In soil quality prediction, temporal feature extraction using GRU facilitated the modeling of temporal relationships among measurements collected over time, making it suitable for applications in which soil quality changed dynamically (Babu *et al.,* 2024). Thus, efficient temporal features were extracted with minimal computational complexity using the proposed ExCSP_Net model.
The output generated from the residual block was provided to the input layer of the GRU. The input sequence was represented as

$$L = \{l_1, l_2, \dots l_f\}$$

where l_f represents the input at time step f. During sequence processing, the GRU computed the hidden state Y_f at each time step. The hidden state estimation was performed through the processing of the reset gate, update gate, and candidate state. The reset gate controlled the retention of relevant information by eliminating unnecessary features and was expressed as:

$$G_f = \varepsilon\left(B_G \cdot \left[Y_{f-1}, l_G\right]\right)$$

where B_G represents the weight associated with the reset gate, l_G is the bias associated with the reset gate, and ε is the sigmoid activation function.
The output of the update gate was calculated as:

$$Z_f = \varepsilon\left(B_Z \cdot \left[Y_{f-1}, l_Z\right]\right)$$

where B_Z is the weight associated with the update gate and l_Z is the bias associated with the update gate.

Subsequently, the candidate state was estimated as:

$$\tilde{Y}_f = tanh\left(B_Y \cdot \left[G_Y \circ Y_{f-1}, l_Y\right] + Q_Y\right)$$

where B_Y is the weight associated with the candidate activation, $\circ$ denotes element-wise multiplication, l_Y is the bias associated with the candidate activation, and Q_Y indicates the additional learnable parameter (Genova *et al.*, 2024).

The hidden state of the GRU was then updated as:

$$Y_f = H_f \circ Y_{f-1} + \left(1 - H_f\right)\tilde{Y}_f$$

where Y_f represents the updated hidden state, H_f is the update control parameter, $Y_{(f-1)}$ is the previous hidden state, and $\tilde{Y}_f$ corresponds to the candidate hidden state.

The GRU allowed the proposed model to capture both short-term and long-term dependencies within soil quality data while maintaining low computational complexity. In addition, it mitigated the vanishing gradient problem and allowed the model to learn from long soil data sequences without losing important information.
Following temporal feature extraction, an attention mechanism was incorporated to assign weights to significant features for improving soil quality forecasting

performance. The attention score for the extracted features from the hidden layer of the GRU was calculated as:

$$p_f = q \cdot tanh\left(B_Y Y_f + Q_Y\right)$$

where q is the learnable vector, B_Y and Q_Y denote the weight and bias parameters, respectively, and p_f represent the attention score.

After computing the attention scores, the weights γ_f were assigned using the SoftMax function and expressed as:

$$\gamma_f = \frac{exp\left(p_f\right)}{\sum_{d=1}^{F} exp\left(p_d\right)}$$

The feature vector generated from the main path of the proposed ExCSP_Net model was defined as:

$$n = \sum_{f=1}^{F} \gamma_f Y_f$$

where n is the combined feature representation obtained from all input sequences. Thus, the output of the main path of the proposed ExCSP_Net model consisted of the most relevant features extracted with minimal computational complexity.

The short path of the ExCSP_Net model utilized the output of the stacked autoencoder for further processing to extract significant features with reduced dimensionality. The encoder compressed the input data into a lower-dimensional representation through multiple fully connected layers. The output of each encoder layer was expressed as:

$$K(r) = R\left(S(r) \cdot K(r-1) + v(r)\right)$$

where $K(r)$ is the output of the rth encoder layer, $S(r)$ denotes the weight matrix, $v(r)$ represents the bias term, R indicates the activation function, and $K(r-1)$ corresponds to the output of the previous layer.

At the center of the stacked autoencoder, a latent representation containing the most salient soil features was generated as:

$$x = R\left(S(x) \cdot K(r) + v(x)\right)$$

where x represents the latent feature vector, while $S(x)$ and $v(x)$ are the associated weights and bias parameters, respectively.

The decoder reconstructed the input data from the latent representation while preserving the significant features of the original input. The decoder output was expressed as:

$$K'(r) = R\big(S(r) \cdot K'(r-1) + v(r)\big)$$

where $K'(r)$ represents the output of the rth decoder layer, and $K'(r-1)$ denotes the output of the previous decoder layer.

The dimensionality-reduced features obtained using the stacked autoencoder were subsequently forwarded through the short path of the ExCSP_Net model. Cross-stage fusion was then performed on the features generated from both the main and short paths of the proposed ExCSP_Net model. The cross-stage fusion mechanism ensured that gradient information from both paths contributed to weight updating while reducing redundant gradients. The output of the fusion layer was expressed as:

$$E_{fused} = Concat(E_M, E_S)$$

where E_{fused} is the merged feature representation, and $Concat(\cdot)$ is the concatenation function.

Subsequently, pyramid pooling was performed to process multi-scale feature representations. The pyramid pooling operation was formulated as:

$$E_p = Concat\left(Pool(E_{fused}, y_1), Pool(E_{fused}, y_2), \ldots Pool(E_{fused}, y_z)\right)$$

where E_p represents the output of the pyramid pooling layer, and $y_1, y_2, \ldots, y_z$ are different pooling kernel sizes used for multi-scale feature extraction.
After the pooling operation, a fully connected network was used for soil quality prediction. The fully connected network consisted of multiple convolution layers, and its output corresponded to the predicted soil quality based on the input soil data. The output of the fully connected layer was formulated as:

$$Out = Conv(E_p, S)$$

where Out represents the predicted output, E_p is the output obtained from pyramid pooling, $Conv(\cdot)$ is the convolution operation, and S indicates the weight parameters. Thus, the proposed ExCSP_Net model enabled more accurate soil quality prediction.

RESULTS AND DISCUSSION

The pre-processing techniques, including missing data imputation and normalization, improved the quality of the input data by reducing noise and inconsistencies, contributing to higher prediction accuracy (Figure 3). The AdPo algorithm ensured that only the most relevant features were selected, which improved the decision-making capability of the model and further enhanced prediction accuracy. In addition, the ExCSP_Net architecture, incorporating GRU-based attention and stacked autoencoders, captured complex patterns in soil data more effectively, improving the overall prediction performance.

The stacked autoencoder incorporated within the ExCSP_Net model performed dimensionality reduction without losing critical information, minimizing overall prediction errors and reducing MAE and RMSE values. In addition, the combination

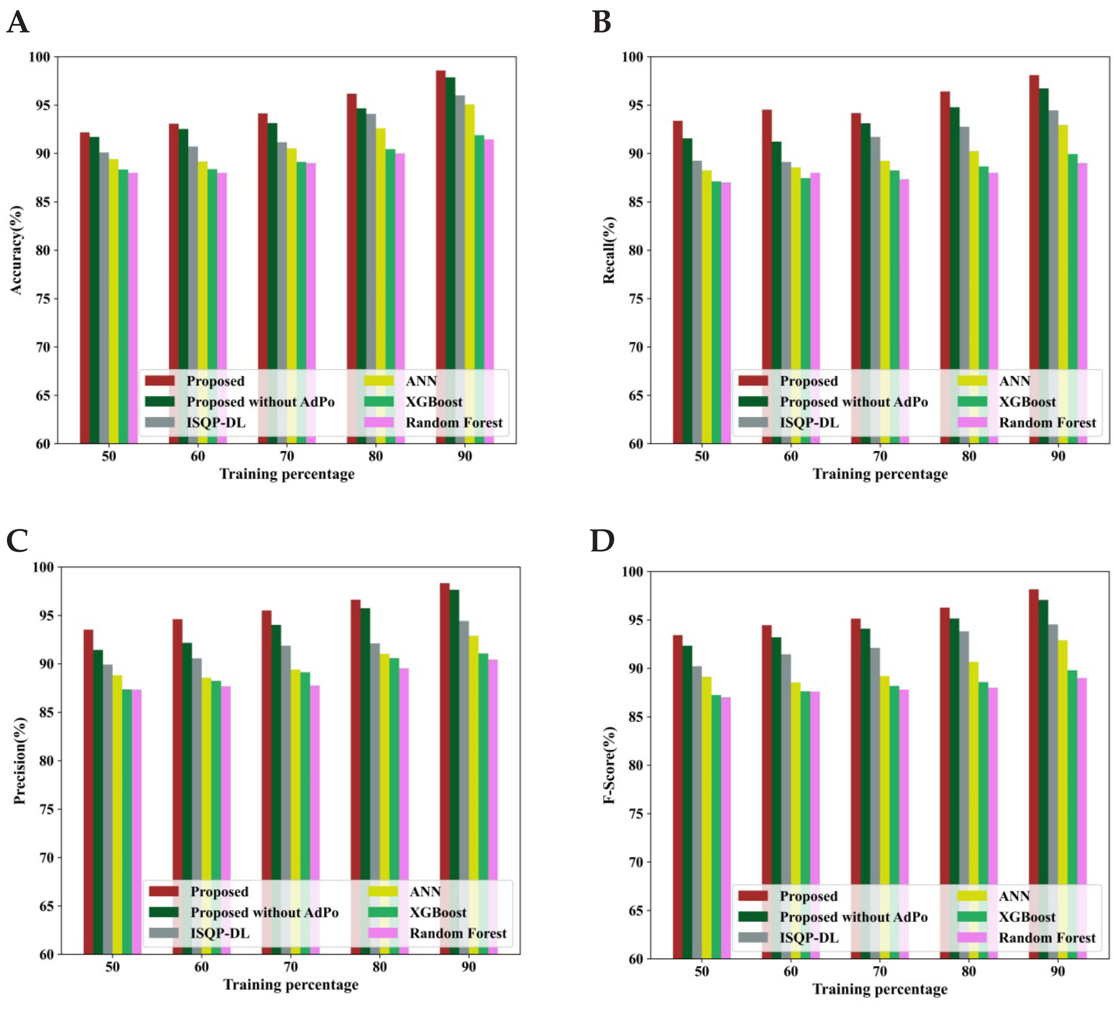

Figure 3. Performance evaluation of soil quality forecasting models based on A: accuracy; B: recall; C: precision; D: F-score.

of AdPo-based feature selection and ExCSP_Net-based soil quality forecasting enabled the model to capture complex relationships between input features and soil quality more effectively, resulting in a higher R^2 value (Figure 4).

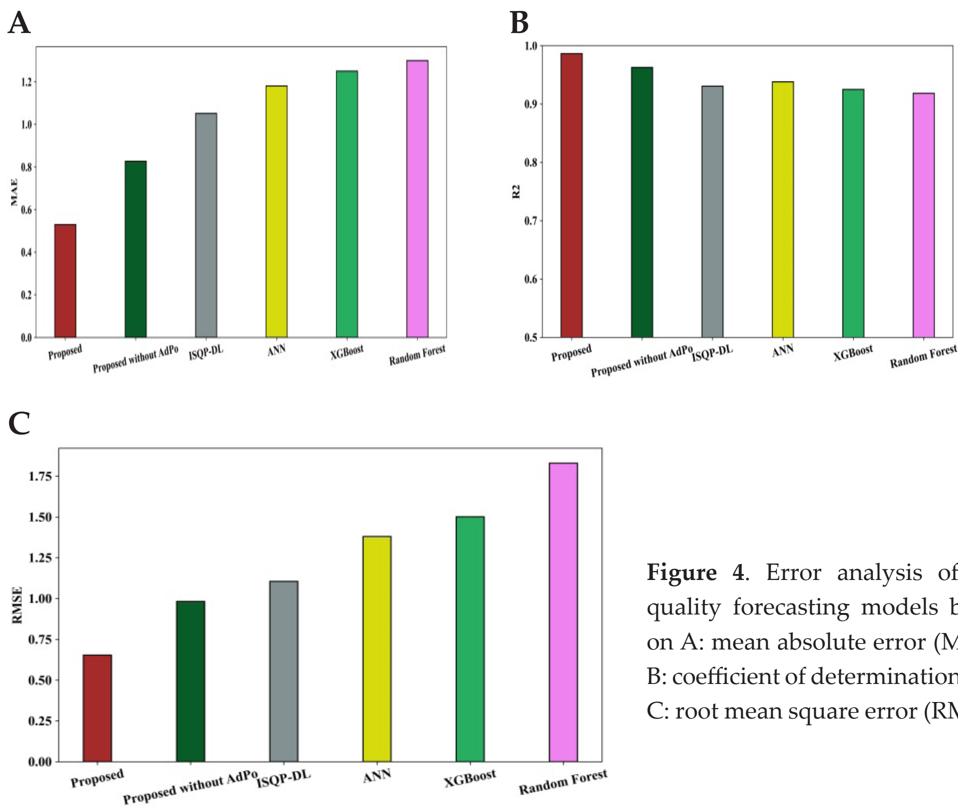

Figure 4. Error analysis of soil quality forecasting models based on A: mean absolute error (MAE); B: coefficient of determination (R^2); C: root mean square error (RMSE).

K-fold validation was performed to demonstrate the generalization capability of the proposed AdPo + ExCSP_Net model (Figure 5). In this approach, the dataset was partitioned into K mutually exclusive and equally sized folds. During each iteration, K−1 folds were used for training the model, while the remaining fold was utilized for validation. The proposed model demonstrated improved performance across all K-folds.

Confusion matrix analysis was performed to quantitatively evaluate the predictive performance of the proposed soil quality forecasting model (Figure 6). The confusion matrix metrics reflected the capability of the model to accurately identify soil quality classes while minimizing incorrect predictions, validating the effectiveness of the AdPo-based feature selection approach and the ExCSP_Net architecture.

SHapley Additive exPlanations (SHAP) analysis was performed (Figure 7) to demonstrate the interpretability of the model by quantifying the contribution of each soil feature selected by the AdPo algorithm to soil quality prediction.

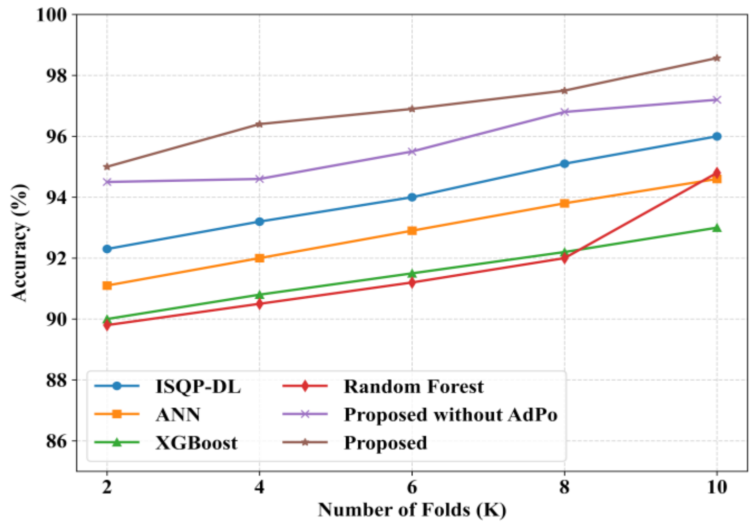

Figure 5. K-fold cross-validation performance of the proposed Adaptive Parrot Optimization with Extended Cross Stage Pyramid Network (AdPo + ExCSP_Net) soil quality forecasting model.

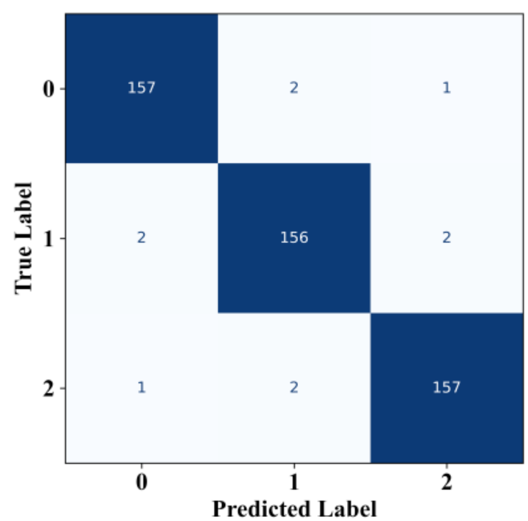

Figure 6. Confusion matrix of the proposed Adaptive Parrot Optimization with Extended Cross Stage Pyramid Network (AdPo+ExCSP_Net) model for soil quality forecasting.

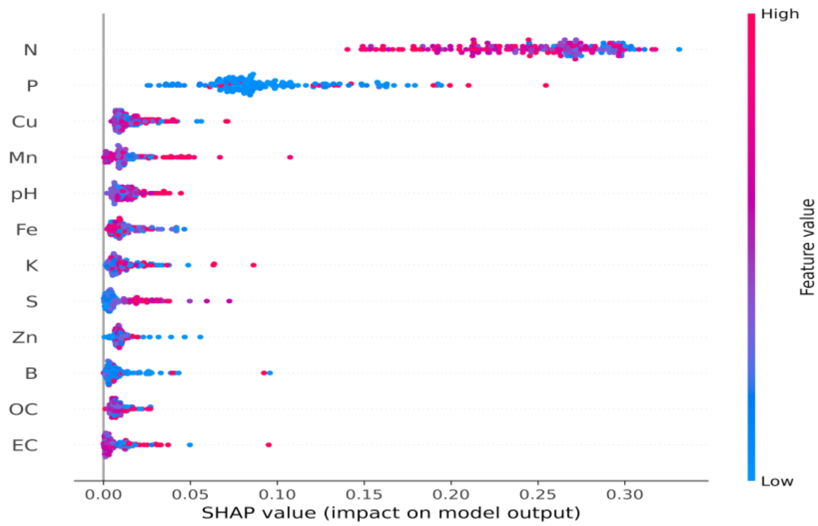

Figure 7. SHapley Additive exPlanations (SHAP)-based analysis of feature contributions in the proposed Adaptive Parrot Optimization with Extended Cross Stage Pyramid Network (AdPo+ExCSP_Net) model.

The proposed AdPo+ExCSP_Net model achieved an accuracy of 98.58 %, which was 2.61, 3.56, 6.78, 7.23, and 0.72 % higher than those obtained using the ISQP-DL, ANN, XGBoost, Random Forest, and proposed model without AdPo methods, respectively (Table 1). Similarly, the proposed method demonstrated improved performance across all evaluation metrics. Thus, the analysis portrays the superiority of the proposed AdPo-based feature selection and ExCSP_Net-based soil quality forecasting.

Table 1. Comparative performance evaluation of soil quality forecasting models based on the best obtained results.

Metrics/ Methods	ISQP-DL	ANN	XGBoost	Random Forest	Proposedmodel without AdPo	Proposed model
Accuracy	96.01	95.07	91.89	91.45	97.87	98.58
Recall	94.45	92.95	89.94	89.00	96.71	98.09
Precision	94.44	92.89	91.07	90.43	97.65	98.32
F-score	94.53	92.89	89.79	89.00	97.07	98.15
MAE	1.05	1.18	1.25	1.30	0.83	0.53
RMSE	1.11	1.38	1.50	1.83	0.98	0.65
R^2	0.93	0.94	0.93	0.92	0.96	0.99

ISQP-DL: improved soil quality prediction-deep learning; ANN: artificial neural network; XGBoost: extreme gradient boosting; MAE: mean absolute error; RMSE: root mean square error; R^2: coefficient of determination.

CONCLUSIONS

This research introduced a new model for forecasting soil quality based on optimal feature selection and deep learning. The proposed Adaptive Parrot Optimization (AdPo) algorithm was designed to perform optimal feature selection from the preprocessed data. The AdPo algorithm enhanced the convergence rate for identifying the global optimal solution through the incorporation of an adaptive weighting strategy, obtaining efficient and non-redundant features.

Soil quality forecasting was performed using the proposed Extended Cross Stage Pyramid Network (ExCSP_Net) model. The developed framework improved forecasting performance through enhanced short and main paths capable of extracting spatial-temporal features with reduced computational complexity. The integration of AdPo with ExCSP_Net demonstrated the effectiveness of combining optimal feature selection with deep learning for accurate soil quality forecasting and sustainable agricultural management. Despite the promising performance of the proposed framework, the computational burden remains a limitation. Therefore, future research will focus on the development of a lightweight model to further improve computational efficiency and practical applicability.

REFERENCES

Babu CS, Yadavamuthiah K. 2024. Soil quality prediction using deep learning. *In* Kumar R, Abdul Hamid AB, Ya'akub DNIB, Ya'akub NIB, Sharan HO, Kumar S. (eds.), Sustainable Development in AI, Blockchain, and E-Governance Applications. IGI Global Scientific Publishing: Hershey, PA, USA, pp: 171–188. https://doi.org/10.4018/979-8-3693-1722-8.ch010

Barathkumar S, Sellamuthu KM, Sathyabama K, Malathi P, Kumaraperumal R, Devagi P. 2025. Advancements in soil quality assessment: A comprehensive review of machine learning and AI-Driven approaches for nutrient deficiency Analysis. Communications in Soil Science and Plant Analysis 56 (2): 251–276. https://doi.org/10.1080/00103624.2024.2406484

Chaudhry H, Vasava HB, Chen S, Saurette D, Beri A, Gillespie A, Biswas A. 2024. Evaluating the soil quality index using three methods to assess soil fertility. Sensors 24 (3): 864. https://doi.org/10.3390/s24030864

Chen X, Zhang X, Wei Y, Zhang S, Cai C, Guo Z, Wang J. 2023. Assessment of soil quality in a heavily fragmented micro-landscape induced by gully erosion. Geoderma 431: 116369. https://doi.org/10.1016/j.geoderma.2023.116369

Divya A, Josphineleela R, Jaba Sheela L. 2024. A machine learning based approach for prediction and interpretation of soil properties from soil spectral data. Journal of Environmental Biology 45 (1): 96–105. https://doi.org/10.22438/jeb/45/1/mrn-5134

El Behairy RA, El Arwash HM, El Baroudy AA, Ibrahim MM, Mohamed ES, Rebouh NY, Shokr MS. 2024. An accurate approach for predicting soil quality based on machine learning in drylands. Agriculture 14 (4): 627. https://doi.org/10.3390/agriculture14040627

Genova G, Borruso L, Signorini M, Mitterer M, Niedrist G, Cesco S, Felderer B, Cavani L, Mimmo T. 2024. Analyzing soil enzymes to assess soil quality parameters in long-term copper accumulation through a machine learning approach. Applied Soil Ecology 195: 105261. https://doi.org/10.1016/j.apsoil.2023.105261

Ghani S, Thapa I, Adhikari DK, Waris KA. 2024. Soil categorization and liquefaction prediction using deep learning and ensemble learning algorithms. Transportation Infrastructure Geotechnology 12 (1). https://doi.org/10.1007/s40515-024-00474-8

Huang Y, Srivastava R, Ngo C, Gao J, Wu J, Chiao S. 2023. Data-driven soil analysis and evaluation for smart farming using machine learning approaches. Agriculture 13 (9): 1777. https://doi.org/10.3390/agriculture13091777

Kalyani NL, Kolla BP. 2024. Machine learning-based smart agricultural practices to assess soil fertility and nutrient dynamics. South Eastern European Journal of Public Health 24 (4): 1363–1378. https://doi.org/10.70135/seejph.vi.2873

Li Y, Yao G, Li S, Dong X. 2025. Predicting and mapping of soil organic matter with machine learning in the black soil region of the southern northeast plain of China. Agronomy 15 (3): 533. https://doi.org/10.3390/agronomy15030533

Omondiagbe OP, Lilburne L, Licorish SA, MacDonell SG. 2023. Soil texture prediction with automated deep convolutional neural networks and population-based learning. Geoderma 436: 116521. https://doi.org/10.1016/j.geoderma.2023.116521

Oukhattar M, Gadal S, Robert Y, Saby N, Houmma IH, Keller C. 2025. Variability analysis of soil organic carbon content across land use types and its digital mapping using machine learning and deep learning algorithms. Environmental Monitoring and Assessment 197 (5). https://doi.org/10.1007/s10661-025-13972-0

Pant J, Joshi MC, Singh D, Pant HK, Bhatt A, Pant D. 2024. An investigation into ensemble learning techniques for evaluating soil fertility through analytical approaches. Procedia Computer Science 235: 1998–2008. https://doi.org/10.1016/j.procs.2024.04.189

Parewai I, Köppen M. 2025. A digital twin approach for soil moisture measurement with physically based rendering simulations and machine learning. Electronics 14 (2): 395. https://doi.org/10.3390/electronics14020395

Raza I, Zubair M, Zaib M, Khalil MH, Haidar A, Sikandar A, Ashfaq MA. 2023. Precision nutrient application techniques to improve soil fertility and crop yield: A review with future prospect. International Research Journal of Educational and Technology 5 (8): 105–123.

Safaie M, Hosseinpour-Zarnaq M, Omid M, Sarmadian F, Ghasemi-Mobtaker H. 2023. Using deep neural networks for evaluation of soil quality based on VIS-NIR spectroscopy. Earth Science Informatics 17 (1): 271–281. https://doi.org/10.1007/s12145-023-01168-4

Selvanarayanan R, Rajendran S, Algburi S, Ibrahim Khalaf O, Hamam H. 2024. Empowering coffee farming using counterfactual recommendation based RNN driven IoT integrated soil quality command system. Scientific Reports 14 (1). https://doi.org/10.1038/s41598-024-56954-x

Shahare YR, Singh MP, Singh SP, Singh P, Diwakar M. 2024. ASUR: Agriculture soil fertility assessment using random forest classifier and regressor. Procedia Computer Science 235: 1732–1741. https://doi.org/10.1016/j.procs.2024.04.164

Shahzad A, Ibrahim M, Saleem RM, Zia S, Habib S, Mahmood M. 2024. Unveiling soil fertility patterns via image analysis and machine learning for accurate crop recommendations. Journal of Computing and Biomedical Informatics.

Song J, Shi X, Wang H, Lv X, Zhang W, Wang J, Li T, Li W. 2024. Improving soil quality index prediction by fusion of Vis-NIR and pXRF spectral data. Geoderma 447: 116938. https://doi.org/10.1016/j.geoderma.2024.116938

Sumathi P, Karthikeyan V, Kavitha M, Karthik S. 2023. Improved soil quality prediction model using deep learning for smart agriculture systems. Computer Systems Science and Engineering 45 (2): 1545–1559. https://doi.org/10.32604/csse.2023.027580

Suwardi, Sutiarso L, Wirianata H, Nugroho AP, Sukarman, Primananda S, Moch D, Hariadi B. 2023. Optimization of a soil type prediction method based on the deep learning model and vegetation characteristics. Plant Science Today 11 (1). https://doi.org/10.14719/pst.2926

Ziyadullaev D, Muhamediyeva D, Khujamkulova K, Abdurakhimov D, Maksumkhanova A, Ziyodullaeva G. 2024. Ensemble data mining methods for assessing soil fertility. E3S Web of Conferences 494: 02013. https://doi.org/10.1051/e3sconf/202449402013