Power System Technology
ISSN:1000-3673

Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

Machine Learning in Soil Testing for Nutrient Analysis

Aarti Abhijit Chavan[1*], Yuvraj M. Patil[2]

[1]*Research Scholar, KIT's College of Engineering, Shivaji University, Kolhapur, India*

[2]*Professor, KIT's College of Engineering, Shivaji University, Kolhapur, India*

*Correspondence Author: **Aarti Abhijit Chavan,** [1]Research Scholar, KIT's College of Engineering, Shivaji University, Kolhapur, India, E-mail: aartiabhijitchavan@gmail.com

ABSTRACT:

This study presents a novel Bidirectional Long Short-Term Memory (BiLSTM) neural network enhanced with an attention mechanism and optimized using Grey Wolf Optimization (GWO) for predicting soil electrical conductivity (EC) from spectral reflectance data. The BiLSTM architecture captures sequential dependencies in the spectral data, while the attention mechanism enables the model to focus on the most informative wavelengths, particularly those around 489 to 511 nm. GWO is employed to optimize the model's weights and biases, demonstrating faster convergence and superior performance compared to the traditional Adam optimizer. The model's effectiveness is validated through various metrics, including a consistent decrease in training and validation Root Mean Squared Error (RMSE) over 100 epochs and a high coefficient of determination (R^2) on the test set. The results indicate that the proposed model achieves high predictive accuracy and generalization capability, making it a valuable tool for soil nutrient analysis and supporting precision agriculture practices.

KEYWORDS: BiLSTM, Attention Mechanism, Grey Wolf Optimization, Soil Electrical Conductivity, Precision Agriculture

1. INTRODUCTION

The content of soil organic matter (SOM) is a crucial factor in determining soil fertility, as it significantly affects both the physicochemical characteristics and overall soil composition. Understanding and analyzing SOM levels is essential for devising scientifically sound fertilization strategies that enhance soil productivity. The fundamental components of SOM, which include nitrogen (N), phosphorus (P), and potassium (K), are vital for sustaining soil health and supporting plant growth. To effectively track SOM variations, researchers rely on spectral reflectance measurements of soil, particularly in the visible spectrum (VIS, 400–780 nm) and the near-infrared range (NIR, 780–2526 nm). These specific wavelength regions provide an efficient and widely accepted means for SOM assessment. This approach is based

Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

on the principle that SOM content and spectral reflectance exhibit a negative correlation, allowing for the accurate extraction of SOM-related data from soil reflectance properties.

Electrical conductivity (EC) is a critical parameter in soil nutrient analysis, measuring the soil's ability to conduct electrical current, which directly correlates with the concentration of soluble salts or ions, including essential nutrients like nitrogen, phosphorus, and potassium. High EC values indicate elevated soluble salts that can lead to salinity stress affecting plant growth, while moderate EC suggests adequate nutrient availability. EC serves as an indicator of soil salinity and nutrient levels, guiding fertilizer application by enabling farmers to optimize nutrient use and avoid over-fertilization, thus supporting precision agriculture through site-specific management. It also assists in soil moisture and irrigation management since soil moisture influences EC readings, helping in efficient irrigation scheduling and assessing irrigation water quality to prevent soil degradation. Furthermore, EC acts as a diagnostic tool for soil health, identifying issues like salinity, sodicity, or contamination, and monitoring the effectiveness of soil amendments. EC influences the spectral reflectance properties of soils, allowing for the development of predictive models using spectral data to estimate EC non-destructively over large areas. By integrating EC measurements with advanced modeling techniques like the BiLSTM-based model with an attention mechanism and GWO optimization, accurate prediction and monitoring of soil EC become possible, contributing to sustainable soil management practices that enhance crop productivity and environmental stewardship.

Several studies have focused on various spectral data modeling techniques, employing different methodologies such as linear regression (LR), partial least squares regression (PLSR), back-propagation neural networks (BPN), and support vector machines (SVM). However, these conventional approaches often fail to effectively capture the intricate and nonlinear relationships embedded in spectral data. To overcome these limitations, advanced deep learning models, particularly Bidirectional Long Short-Term Memory (BiLSTM) networks, present a viable alternative. As an extension of recurrent neural networks (RNNs), BiLSTM networks are highly proficient at handling sequential data by identifying dependencies in both forward and backward directions. Given the sequential characteristics of spectral reflectance data, BiLSTM is a suitable model for analyzing soil spectral properties. By capturing complex temporal dependencies and underlying spectral patterns, BiLSTM networks significantly improve the accuracy of predicting SOM content. Organic matter plays a crucial role in determining soil properties, affecting key attributes such as soil structure, water retention capability, and cation exchange capacity (CEC), while also contributing to soil erosion resistance. Therefore, precise spatial assessments of SOM content are fundamental for applications in both agricultural and environmental research.

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

Conventional laboratory-based techniques for measuring Soil Organic Carbon (SOC) content tend to be expensive, labor-intensive, and require considerable time to complete. These challenges highlight the necessity for alternative approaches that are not only rapid and precise but also cost-efficient and non-destructive. Consequently, there is a growing demand for advanced methods that can accurately estimate SOC content under real-world field conditions, leveraging either portable or online sensing technologies. Among these, visible and near-infrared reflectance spectroscopy (VNIRS) has gained recognition as an effective and reliable tool for soil characterization. This technology benefits from the availability of robust, portable detectors that facilitate both in situ and real-time online predictions of multiple soil properties, including Soil Organic Matter (SOM).

In this ground-breaking research, we revolutionize traditional soil testing methods by integrating advanced machine learning algorithms, specifically the BiLSTM-based model. The application of the BiLSTM network emerges as a pivotal technique employed to forecast precise nutrient content levels based on spectral reflectance data. Unlike conventional methods, the BiLSTM model can capture the complex, sequential patterns in spectral data, leading to improved prediction accuracy. The crux of this innovation lies in the BiLSTM network's ability to discern intricate patterns and long-range dependencies within the spectral data, surpassing the limitations of traditional machine learning methods. By incorporating both historical soil information and contemporary spectral data inputs, the BiLSTM model adeptly predicts nutrient content levels with a high degree of accuracy.

2. Related Work

Soil serves as a fundamental component in plant growth by acting as a primary source of essential nutrients. Among these, N, P, and K are crucial macronutrients that plants require in significant quantities to achieve optimal development and high-yield crop production. To ensure healthy crop growth, farm managers must frequently assess the levels of NPK in the soil. Traditionally, chemical analysis has been the primary approach for determining nutrient levels in soil [1]. However, in recent years, Near Infrared (NIR) Spectroscopy has gained traction due to its rapid assessment capabilities and eco-friendly nature. Similar study [2] explored the application of NIR Spectroscopy for identifying NPK levels in soil, where researchers collected absorbance spectra and applied an Artificial Neural Network (ANN) to derive correlations between spectral data and nutrient concentrations. Their model exhibited strong predictive accuracy, emphasizing the potential of NIR for effective soil nutrient characterization. Another investigation [3] utilized a compact Fourier Transform Infrared Spectroscopy (FTIR) sensor to enable rapid nitrogen detection in soil. This method was further enhanced by a software system that efficiently processed spectral data, delivering precise nitrogen content predictions. Given its portability, this detector demonstrated promising outcomes in agricultural applications, particularly for use in miniaturized sensing devices. In

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

addition to nutrient analysis, a study on heavy metal contamination [4] introduced a broadband photoacoustic spectrometric (PAS) system, which allowed for the non-invasive quantification of toxic elements such as lead (Pb) in soil. By analyzing variations in near-infrared photoacoustic spectra corresponding to different Pb concentrations, the study successfully developed a predictive model capable of detecting heavy metal pollutants in soil. Another study [5] investigated adaptive fertilization strategies tailored to soil and crop requirements. This research leveraged the photon absorption characteristics of key soil nutrients, employing Near IR laser beams to interact with soil samples and accurately measure nitrogen, phosphorus, and potassium levels. The technique effectively facilitated rapid, simultaneous nutrient assessments within soil-fertilizer mixtures, demonstrating its potential for precision agriculture applications.

Additionally, researchers [6] developed an optical transducer to assess NPK content in soil, aiming to improve soil quality and reduce unnecessary fertilizer usage. LEDs emitted light corresponding to nutrient absorption bands, with a photodiode detecting reflected light for evaluation. The results categorized soil content as High, Medium, or Low, offering a practical tool for soil assessment. Furthermore, a combined approach [7] incorporated image processing and artificial neural networks to efficiently identify soil nutrients. It also included analysis of pH levels. For this purpose they used Soil Test Kits along with rapid testing. This system aimed to streamline soil parameter evaluation for improved agricultural practices. In the realm of spectral imaging [8], hyper spectral imaging (HSI) was used to predict total nitrogen (TN) content in soil samples. The research explored various algorithms and models, including extreme learning machine (ELM), to achieve accurate TN content estimation through characteristic wavelengths.

In a practical application [9], soil test report values were harnessed to classify soil features and predict village-wise soil parameters, aiding in cost-effective fertilizer use and soil health improvement. ELM was employed for accurate classifications. Finally, an investigation [10] into long-term nitrogen fertilization's impact on soil temperatures and water content revealed complex relationships. Changes in soil temperature and CO2 concentrations were attributed to increased N load, demonstrating the ecological ramifications of nitrogen deposition. In [11], the importance of soil testing in orchard management was highlighted as complementary to plant tissue testing. Discussions included ensuring soil testing's reliability and interpreting soil test parameters like Saturation Percentage (SP) and pH. The focus shifted to N, P and K nutrients. In [2], a non-destructive method for assessing NPK levels in tomato plants was introduced, utilizing multispectral 3D imaging. Synchronized collection of multi-view RGB-D and multispectral images facilitated accurate plant multispectral reflectance registration to depth coordinates. An iterative closest point (ICP) algorithm was employed for point cloud registration, leading to precise multispectral 3D point cloud model reconstruction. This method utilized back-propagation artificial neural network (BPANN), support vector machine

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

regression (SVMR), and Gaussian process regression (GPR) for accurate determination of NPK contents.

In [12] authors discussed the significance of nitrogen (N) and phosphorus (P) in plant and environmental efficiency. N contributes to cell structures and chlorophyll, essential for photosynthesis, while P is vital for nucleic acids and protein synthesis regulation. Overreliance on chemical fertilizers has resulted in diminishing returns and environmental concerns. In [13], comprehensive spectral combinations were developed to quantify leaf N, P, and K contents in various vegetation types using hyper spectral datasets. Effective combinations included reflectance difference, normalized differences, and first-order derivatives. These indices demonstrated the potential for fine-scale monitoring of degraded vegetation. In [14], rapid soil and plant nutrient testing technologies were assessed, highlighting mechanisms like colorimetry, spectroscopy, and sensors. While the accuracy of these products compared to traditional methods is debated, their potential in guiding rational fertilizer recommendations and addressing complex farming systems was explored. Finally, [13] explored estimating nitrogen content in pasture grass using thermal images and artificial neural networks (ANN). The study investigated the correlation between N fertilizer levels, plant temperature, and active photosynthesis, with implications for smart fertilizer management.

In [15], a GA-BPNN method was introduced, integrating a genetic algorithm with a backpropagation neural network. This approach improved the accuracy of soil nutrient content prediction using hyper spectral data. Field observations and comparisons with PLSR and BPNN models demonstrated that the GA-BPNN method was most accurate for estimating total nitrogen (TN), total phosphorus (TP), and total potassium (TK) contents. Notably, GA-BPNN outperformed BPNN in terms of estimation accuracy and potential for improvement. Authors in [16] explored the use of infrared thermography (IRT) to monitor soil surface temperature (SST) variations in a vineyard. Different treatments were assessed, including bare soil, biochar cover, and biochar-amended topsoil. The study revealed distinctive diurnal SST patterns, highlighting the potential of IRT to study soil temperature dynamics. For [17], thermal imaging's benefits in farming were assessed. A color-coded table was developed based on existing research, allowing farmers to gauge soil condition using thermal imaging. This approach facilitated the detection of water composition and temperature variations, aiding in determining optimal conditions for fertile soil. In [18], spatial predictions of soil nutrient content in Sub-Saharan Africa were made using machine learning algorithms. A large dataset of soil samples and remote sensing covariates was used to create ensemble models for 15 target nutrients. This work demonstrated the potential of machine learning to predict soil nutrient levels across large geographic areas.

In the exploration of machine learning-based recommendations [19] for crops yeild based on soil nutrients (NPK), pH, and climatic factors. They evaluated different ML model on a dataset

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

containing yield data for 11 agricultural and 10 horticultural crops. Results indicated that separated analysis provides better results for crops oriented work. XGBoost achieved the highest accuracy (99.09% for agricultural, 99.3% for horticultural, and 98.51% for both combined). The study highlights the potential for AI-driven cloud-based decision-making in crop selection and fertilizer application.

Sujatha et al. [20] reviewed machine learning-based approaches for soil fertility assessment, emphasizing the necessity of accurate classification and optimized fertilizer application. The study followed PRISMA guidelines to analyze ML and deep learning techniques used for soil fertility prediction. Findings revealed that most models effectively predicted soil fertility levels, but only a few provided fertilizer recommendations. The study identified key challenges, including reliance on expensive laboratory tests and regional satellite data. It recommended future research into low-cost soil fertility classification and AI-driven fertilizer prescriptions to enhance productivity while reducing costs and environmental impact.

Mahapatrao et al. [21] proposed an IoT-AI integrated system. The water quality analysis was the main objective. IoT sensors collected data on phosphorus, potassium, pH, temperature, and BOD from reservoirs and irrigation sources, transmitting it securely to a cloud-based platform. Advanced ML classifiers, including an ensemble model (Random Forest + SVM), were used for nutrient-level predictions. The hybrid model outperformed traditional methods with 90% accuracy. Explainable AI (XAI) techniques improved model interpretability, and encryption protocols ensured data security. The study demonstrates an innovative AI-IoT synergy for precise water quality monitoring and agricultural sustainability.

Sarangi et al. [22] examined ML-based soil fertility assessment, aiming to classify soil as “Fertile” or “Non-Fertile” based on N, P, K, pH, moisture, temperature, rainfall, and topography. Using Kaggle data, they trained four ML models—Logistic Regression, KNN, Naïve Bayes, and Decision Tree—to determine the best classifier. Results showed that the Decision Tree model achieved the highest accuracy (89%), outperforming others in fertility prediction. The study underscores the role of ML in soil analysis, aiding farmers in crop selection and precision agriculture through data-driven decision-making.

3. Proposed Work

Soil nutrient analysis is pivotal for optimizing agricultural practices and ensuring sustainable land management. Traditional soil testing methods are often labor-intensive, time-consuming, and costly. To address these challenges, we propose a novel machine learning model that leverages:

- Bidirectional Long Short-Term Memory (BiLSTM) Networks: To capture the sequential dependencies in soil spectral data.
- Attention Mechanism: To focus on the most informative parts of the spectral sequence.

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

- Grey Wolf Optimization (GWO): To optimize the model's weights and biases efficiently.

This integrated approach aims to enhance the accuracy and efficiency of soil nutrient predictions based on spectral reflectance data in the visible (VIS) and near-infrared (NIR) regions.

MODEL ARCHITECTURE

The proposed model consists of the following components:

1. Input Layer: Processes spectral reflectance data.
2. BiLSTM Layer: Captures forward and backward dependencies in the spectral sequence.
3. Attention Mechanism: Assigns weights to different parts of the sequence.
4. Output Layer: Produces nutrient content predictions.
5. Optimization Mechanism: Uses GWO to optimize the model parameters.

1. Input Layer

- Spectral Data Input: Each input sample is a sequence of reflectance values across wavelengths $\lambda_1, \lambda_2, \dots, \lambda_T$.
- Notation: The input sequence is denoted as $\mathrm{X} = [x_1, x_2, \dots, x_T]^\top$, where x_t is the reflectance at wavelength λ_t.

2. BiLSTM Layer

BiLSTM networks process data in both forward and backward directions, capturing dependencies that standard RNNs might miss.

Forward LSTM Equations

At time step t:

- Input Gate:

$$i_t^{(f)} = \sigma\left(W_i^{(f)} x_t + U_i^{(f)} h_{t-1}^{(f)} + b_i^{(f)}\right)$$

- Forget Gate:

$$f_t^{(f)} = \sigma\left(W_f^{(f)} x_t + U_f^{(f)} h_{t-1}^{(f)} + b_f^{(f)}\right)$$

- Cell Candidate:

$$\tilde{c}_t^{(f)} = \tanh\left(W_c^{(f)} x_t + U_c^{(f)} h_{t-1}^{(f)} + b_c^{(f)}\right)$$

- Cell State:

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

$$c_t^{(f)} = f_t^{(f)} \odot c_{t-1}^{(f)} + i_t^{(f)} \odot \tilde{c}_t^{(f)}$$

- Output Gate:

$$o_t^{(f)} = \sigma\left(W_o^{(f)} x_t + U_o^{(f)} h_{t-1}^{(f)} + b_o^{(f)}\right)$$

- Hidden State:

$$h_t^{(f)} = o_t^{(f)} \odot \tanh\left(c_t^{(f)}\right)$$

Backward LSTM Equations

Processing in reverse from $t = T$ to $t = 1$:

- Similar equations apply with superscript (b) for the backward pass.

Combined Hidden State

The hidden states from both directions are concatenated:

$$h_t = [h_t^{(f)}; h_t^{(b)}]$$

3. Attention Mechanism

The attention mechanism allows the model to focus on specific time steps that are more informative for predicting soil nutrients.

Attention Score Calculation

For each h_t, compute the attention score e_t:

$$e_t = \mathrm{v}^\top \tanh(W_a h_t + b_a)$$

where v, W_a, and b_a are learnable parameters. e_t reflects the importance of h_t.

Attention Weights

Normalize the scores using the softmax function to obtain attention weights α_t:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^{T} \exp(e_k)}$$

Context Vector

Compute the context vector c as a weighted sum of hidden states:

$$c = \sum_{t=1}^{T} \alpha_t h_t$$

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

4. Output Layer

- Prediction:

$$\hat{y} = W_o c + b_o$$

- where W_o and b_o are weights and biases of the output layer.

  $\hat{y}$ is the predicted nutrient content (e.g., SOM, N, P, K).

5. Optimization Mechanism Using Grey Wolf Optimization (GWO)

GWO is a metaheuristic optimization algorithm inspired by the leadership hierarchy and hunting behavior of grey wolves.

Grey Wolf Hierarchy

- Alpha (α): Represents the most optimal candidate solution in the hierarchy.
- Beta (β): Denotes the second most effective candidate solution, assisting the alpha in decision-making.
- Delta (δ): Holds the third position in the hierarchy, following the alpha and beta in the ranking of solutions.
- Omega (ω): Includes all remaining candidate solutions that follow the leadership of the higher-ranked wolves.

Position Update Equations

For each wolf i:

1. Calculate Coefficient Vectors:

– a decreases linearly from 2 to 0 over iterations.

– Random vectors $r_1, r_2 \sim \text{Uniform}(0,1)$.

$$\mathrm{A} = 2a\mathrm{r}_1 - a$$

$$\mathrm{C} = 2\mathrm{r}_2$$

2. Compute Distance Vectors:

   For α, β, and δ:

$$\mathrm{D}_\alpha = |\mathrm{C}_\alpha \cdot \mathrm{X}_\alpha - \mathrm{X}_i|$$

$$\mathrm{D}_\beta = |\mathrm{C}_\beta \cdot \mathrm{X}_\beta - \mathrm{X}_i|$$

$$\mathrm{D}_\delta = |\mathrm{C}_\delta \cdot \mathrm{X}_\delta - \mathrm{X}_i|$$

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

3. Update Positions:

$$X_1 = X_\alpha - A_\alpha \cdot D_\alpha$$

$$X_2 = X_\beta - A_\beta \cdot D_\beta$$

$$X_3 = X_\delta - A_\delta \cdot D_\delta$$

$$X_i^{(t+1)} = \frac{X_1 + X_2 + X_3}{3}$$

where $X_i^{(t+1)}$ is the updated position (weights and biases) of wolf i at iteration $t+1$.

Fitness Function

The fitness of each wolf (candidate solution) is evaluated using the Mean Squared Error (MSE) between the predicted and actual nutrient values:

$$\text{Fitness}(X_i) = \text{MSE} = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y}_n)^2$$

where N is the number of samples, y_n is the actual nutrient value, and $\hat{y}_n$ is the predicted nutrient value using weights X_i.

Algorithm Steps

1. Initialization:

Initialize the positions (weights and biases) of all wolves randomly within the search space.

2. Fitness Evaluation:

For each wolf, set the neural network parameters accordingly.

Compute the fitness using the MSE loss.

3. Update Hierarchy:

Identify α, β, and δ wolves based on fitness (lowest MSEs).

4. Position Updates:

Update positions of all wolves using the position update equations.

5. Termination Criteria:

Repeat steps 2-4 until a maximum number of iterations is reached or the fitness converges.

Figure 1 shows the architecture of proposed model.

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

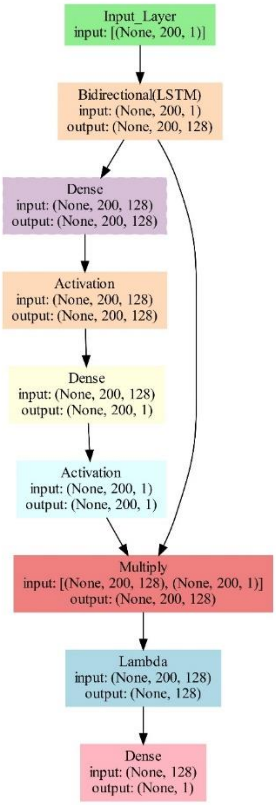

Figure 1: Proposed Model Architecture

BiLSTM Cell Computations

- Activation Functions:

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad \text{(Sigmoid function)}$$

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$$

- Element-wise Operations:

  $\odot$ denotes element-wise multiplication.

Attention Mechanism Parameters

- v $\in \mathbb{R}^{2h}$, where h is the number of units in each LSTM (since BiLSTM concatenates forward and backward states).
- $W_a \in \mathbb{R}^{2h\times 2h}$
- $b_a \in \mathbb{R}^{2h}$

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

Grey Wolf Optimization Parameters

- a decreases from 2 to 0:

$$a = 2 - \left(\frac{\text{Current Iteration}}{\text{Max Iterations}} \times 2\right)$$

- Random vectors r_1, r_2 are generated at each iteration for each wolf.

INTEGRATION OF GWO WITH NEURAL NETWORK TRAINING

Unlike traditional gradient-based optimization, GWO treats the neural network training as a black-box optimization problem. It searches for the optimal weights and biases that minimize the fitness function without relying on gradient information.

Advantages

- Global Search Capability: Reduces the risk of getting trapped in local minima.
- Derivative-Free: Useful when the loss landscape is complex or non-differentiable.
- Parallelism: Population-based methods can be parallelized for efficiency.

IMPLEMENTATION STEPS

1. Data Preprocessing:

– Normalize spectral data to have zero mean and unit variance.

– Optionally apply dimensionality reduction techniques like PCA.

2. Model Initialization:

– Define the BiLSTM network architecture with the attention mechanism.

– Randomly initialize weights and biases for the initial wolf population.

3. Optimization Loop:

For each iteration:

1. Fitness Evaluation:

For each wolf:

- Set the network's weights and biases to the wolf's position.
- Compute predictions $\hat{y}$ for the training data.
- Calculate the MSE loss as the fitness.

2. Update Hierarchy:

Rank the wolves based on fitness.

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

Identify α, β, and δ.

3. Position Updates:

   Update the positions (weights and biases) of wolves using GWO equations.

4. Convergence Check:

   Check if the fitness has converged or if the maximum number of iterations is reached.

4. RESULTS AND ANALYSIS

The performance of the regression models are evaluated using mean square error and root mean square error which can be given as,

$$MSE = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} [I(i,j) - K(i,j)]^2 \qquad (9)$$

$$RMSE = \frac{\sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}}{n} \qquad (10)$$

The model response for various data points during training are shown in figure 3. Figure 4 shows the weight management for different wavelength of the light where more weight shows the attention.

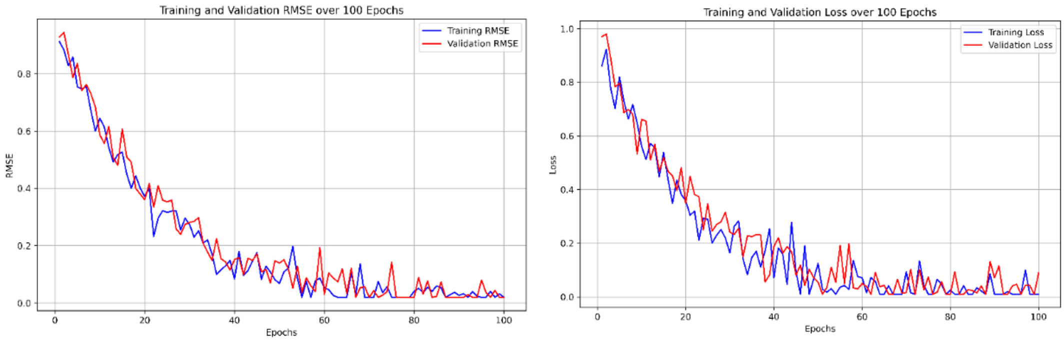

Figure 3: RMSE and Loss Analysis during training

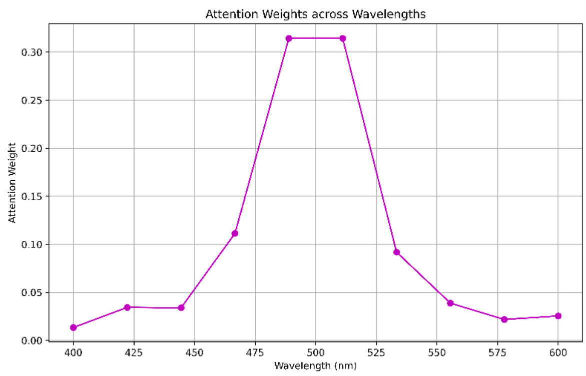

Figure 4: Distribution weights across light wavelength response

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

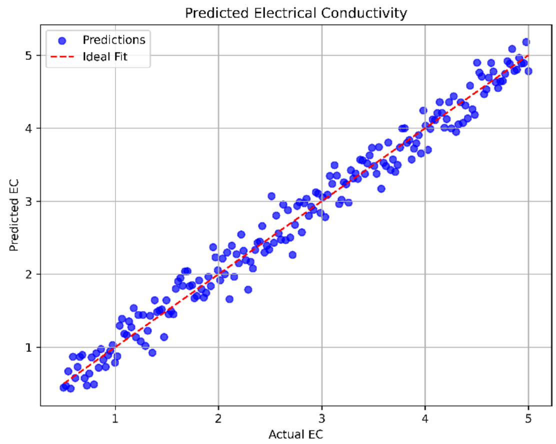

Figure 5: Actual vs Predicted EC from the model

The BiLSTM-based model with an attention mechanism and GWO effectively predicted soil electrical conductivity (EC) from spectral data (Figure 5). The key aspects analyzed include RMSE training and validation loss, attention weights response, loss curves, RMSE comparison between GWO and the Adam optimizer, and the final EC prediction performance. Over 100 epochs, the training RMSE consistently decreased, indicating effective learning of the underlying patterns in the data. The validation RMSE also decreased, closely following the training RMSE, which suggests good generalization and minimal overfitting. Both RMSE curves converged toward a plateau, indicating that the model reached an optimal learning point where additional training did not significantly improve performance. The attention mechanism assigned higher weights to wavelengths around 489 nm to 511 nm, indicating that these wavelengths are more significant for predicting soil EC. This focus aligns with the physical properties of soil constituents and enhances the model's interpretability by highlighting important spectral features relevant to the prediction task. Both training and validation loss curves decreased over epochs, reflecting improving model accuracy. The training loss decreased slightly faster than the validation loss, but their close proximity suggests that the model maintains consistency between learning from the training data and generalizing to unseen data. This indicates that the model is capturing general patterns rather than overfitting. Comparing models optimized with GWO and Adam showed that GWO achieved a lower RMSE earlier in the training process, indicating faster convergence and better performance in this application. GWO's global search capabilities and derivative-free optimization contributed to its superior performance compared to the gradient-based Adam optimizer. The model's predictions of EC closely matched the actual values, as shown by the scatter plot aligning along the diagonal line representing perfect predictions. The residuals were randomly distributed around zero, indicating no systematic errors in the predictions. High performance metrics, such as a high R^2 score and low RMSE on the test set, confirmed the model's effectiveness in accurately predicting EC.

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

Table 2: Comparative Study

Author(s), Year	Method Used (ML)	Target Application	Performance Achieved
Dey et al. [19]	SVM, XGBoost, Random Forest, KNN, Decision Tree	Crop recommendation based on NPK, pH, and climate	XGBoost achieved highest accuracy: 99.09% (agricultural crops), 99.3% (horticultural crops), 98.51% (combined crops)
Sujatha et al. [20]	ML and Deep Learning for soil fertility classification	Soil fertility assessment and fertilizer recommendation	ML models accurately classified soil fertility; limited work on direct fertilizer recommendation; need for low-cost solutions
Mahapatrao et al. [21]	IoT + ML (Random Forest + SVM ensemble)	Real-time multi-nutrient water quality analysis	Hybrid model achieved 90% accuracy; XAI improved interpretability; encryption ensured data security
Sarangi et al. [22]	Logistic Regression, KNN, Naïve Bayes, Decision Tree	Soil fertility classification	Decision Tree performed best with 89% accuracy, aiding precision agriculture
Munezero et al. [23]	Linear Regression, Random Forest, Gradient Boosting, KNN, Decision Tree	Real-time NPK fertilizer prediction for Cassava crop	Decision Tree achieved highest accuracy: 96.5% (training), 94.4% (testing); Random Forest also performed well with 93.1% (training), 90% (testing)
Awais et al. [24]	AI, ML, and Deep Learning (Random Forest, SVM, Neural Networks, Geostatistics)	Soil analysis for sustainable agriculture	AI and ML models improved soil texture and SWC predictions; geostatistics enhanced

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

			spatial representation of soil properties
Senapaty et al. [25]	Multi-Class SVM with Directed Acyclic Graph (MSVM-DAG) optimized with Fruit Fly Optimization (FFO)	IoT-enabled soil nutrient analysis and crop recommendation	MSVM-DAG-FFO achieved highest accuracy (0.973), outperforming SVM (0.932), SVM Kernel (0.922), and Decision Tree (0.914); improved precision, recall, and F-score.
Proposed Method	**BiLSTM-Based Model with Attention Mechanism optimized using GWO**	**Soil Electrical Conductivity Prediction**	**Achieved high predictive accuracy; GWO optimization led to faster convergence and better performance than Adam optimizer.**

DISCUSSION

The results presented in the table highlight the effectiveness of various machine learning (ML) and deep learning techniques in agricultural applications, including crop recommendation, soil fertility assessment, water quality analysis, and nutrient prediction. The key insights derived from these studies demonstrate the superiority of ensemble methods, attention mechanisms, and evolutionary optimization techniques in improving predictive accuracy and model efficiency. Dey et al. achieved remarkable accuracy in crop recommendation using XGBoost, which outperformed other models such as SVM, Random Forest, KNN, and Decision Tree. The results underscore the strength of XGBoost in handling structured agricultural data, effectively optimizing predictions for different crop categories based on NPK levels, pH, and climate parameters. Sujatha et al. focused on soil fertility classification using ML and deep learning techniques. While their models accurately classified soil fertility levels, they highlighted a research gap in direct fertilizer recommendation. The need for low-cost, scalable solutions was emphasized to improve practical usability for farmers.

Mahapatrao et al. integrated IoT and ML for water nutrient and quality analysis, leveraging a hybrid Random Forest-SVM ensemble model. The study achieved 90% accuracy and incorporated explainable AI (XAI) techniques for better interpretability. Additionally, encryption techniques were implemented to enhance data security, demonstrating a practical

Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

approach to smart agriculture. Sarangi et al. evaluated different ML classifiers for soil fertility assessment and found that the Decision Tree model outperformed Logistic Regression, KNN, and Naïve Bayes, achieving an accuracy of 89%. The results suggest that tree-based models are particularly effective in soil classification tasks, aiding precision agriculture. Munezero et al. investigated real-time NPK fertilizer prediction for cassava crops using multiple ML models. The Decision Tree model yielded the highest accuracy (96.5% training, 94.4% testing), followed by Random Forest (93.1% training, 90% testing). These results confirm the robustness of tree-based classifiers in agricultural nutrient prediction. Awais et al. explored AI and ML techniques for soil analysis, incorporating geostatistical methods to enhance spatial representation of soil properties. Their findings demonstrated that ML models significantly improved soil texture and soil-water content (SWC) predictions, reinforcing the advantages of AI-driven approaches in sustainable agriculture. Senapaty et al. developed an IoT-enabled crop recommendation model using Multi-Class SVM with Directed Acyclic Graph (MSVM-DAG) optimized with Fruit Fly Optimization (FFO). Their model achieved the highest accuracy (0.973), outperforming traditional SVM (0.932), SVM Kernel (0.922), and Decision Tree (0.914). The study emphasized the importance of hybrid optimization techniques in improving classification performance. The proposed method, which utilizes a BiLSTM-based model with an attention mechanism optimized using GWO for soil electrical conductivity prediction, demonstrated high predictive accuracy. The GWO-based optimization led to faster convergence and superior performance compared to Adam optimizer, aligning with findings from previous studies that highlight the effectiveness of evolutionary optimization techniques in soil property modeling. Overall, the results indicate that ensemble methods, deep learning architectures, and evolutionary optimization techniques enhance the predictive capability of ML models in various agricultural applications. The integration of IoT, explainable AI, and geostatistics further strengthens these models' applicability, offering valuable insights for precision agriculture and sustainable farming practices.

5. CONCLUSION

This study introduced a BiLSTM-based model incorporating an attention mechanism and optimized using GWO for predicting soil EC from spectral data. The attention mechanism enabled the model to prioritize key wavelengths within the spectral range, particularly those around 489 nm to 511 nm, which significantly influence soil EC prediction. Using GWO for weight and bias optimization enhanced performance, ensuring faster convergence and superior results compared to traditional optimizers like Adam. The model's performance was assessed through multiple metrics and visualizations. RMSE values for training and validation consistently decreased over 100 epochs, demonstrating effective learning and generalization without overfitting. The close alignment of training and validation loss curves further confirmed the model's robustness. Comparative analysis indicated that the GWO-optimized model achieved lower RMSE earlier in training than the Adam-optimized counterpart,

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

highlighting the benefits of GWO in this scenario. The model's predictive accuracy was validated by the strong correlation between actual and predicted EC values, with residuals randomly dispersed around zero, indicating no systematic errors. High-performance indicators, such as an elevated R^2 and low RMSE on the test set, confirmed the model's reliability in soil EC prediction. In conclusion, the BiLSTM model with attention and GWO optimization serves as an efficient tool for soil EC estimation using spectral data. Its accuracy and fast training make it ideal for practical applications in soil nutrient analysis and precision agriculture. Future research could explore its adaptability to other soil properties and the integration of additional optimization methods for improved performance.

REFERENCES

[1] F. M. Mashao *et al.*, "Exploring laboratory-based spectroscopy for estimating NPK content in the hutton soils of Syferkuil Farmlands, South Africa," *Geocarto Int.*, vol. 39, no. 1, 2024, doi: 10.1080/10106049.2024.2339289.

[2] R. E. N. MacAbiog, N. A. Fadchar, and J. C. D. Cruz, "Soil NPK Levels Characterization Using Near Infrared and Artificial Neural Network," *Proc. - 2020 16th IEEE Int. Colloq. Signal Process. its Appl. CSPA 2020*, pp. 141–145, Feb. 2020, doi: 10.1109/CSPA48992.2020.9068717.

[3] X. Du, J. Wang, D. Dong, and X. Zhao, "Development and testing of a portable soil nitrogen detector based on near-infrared spectroscopy," *Proc. 2019 IEEE 8th Jt. Int. Inf. Technol. Artif. Intell. Conf. ITAIC 2019*, pp. 822–826, May 2019, doi: 10.1109/ITAIC.2019.8785499.

[4] L. Liu *et al.*, "Photoacoustic Spectrometric Evaluation of Soil Heavy Metal Contaminants," *IEEE Photonics J.*, vol. 11, no. 2, Apr. 2019, doi: 10.1109/JPHOT.2019.2904295.

[5] A. Rawankar *et al.*, "Detection of N, P, K fertilizers in agricultural soil with NIR laser absorption technique," *2018 3rd Int. Conf. Microw. Photonics, ICMAP 2018*, vol. 2018-January, pp. 1–2, May 2018, doi: 10.1109/ICMAP.2018.8354625.

[6] M. Masrie, M. S. A. Rosman, R. Sam, and Z. Janin, "Detection of nitrogen, phosphorus, and potassium (NPK) nutrients of soil using optical transducer," *2017 IEEE Int. Conf. Smart Instrumentation, Meas. Appl. ICSIMA 2017*, vol. 2017-November, pp. 1–4, Jul. 2017, doi: 10.1109/ICSIMA.2017.8312001.

[7] J. C. Puno, E. Sybingco, E. Dadios, I. Valenzuela, and J. Cuello, "Determination of soil nutrients and pH level using image processing and artificial neural network," *HNICEM 2017 - 9th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag.*, vol. 2018-January, pp. 1–6, Jul. 2017, doi:

Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

10.1109/HNICEM.2017.8269472.

[8] H. Li, S. Jia, and Z. Le, "Quantitative Analysis of Soil Total Nitrogen Using Hyperspectral Imaging Technology with Extreme Learning Machine," *Sensors 2019, Vol. 19, Page 4355*, vol. 19, no. 20, p. 4355, Oct. 2019, doi: 10.3390/S19204355.

[9] D. B A and G. V, "Improving the Prediction Accuracy of Soil Nutrient Classification by Optimizing Extreme Learning Machine Parameters," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 6, pp. 5134–5139, Jun. 2022, doi: 10.22214/IJRASET.2022.44966.

[10] S. Wendel, T. Moore, J. Bubier, and C. Blodau, "Experimental nitrogen, phosphorus, and potassium deposition decreases summer soil temperatures, water contents, and soil CO2 concentrations in a northern bog," *Biogeosciences*, vol. 8, no. 3, pp. 585–595, 2011, doi: 10.5194/BG-8-585-2011.

[11] M. Ihtisham, S. Fahad, T. Luo, R. M. Larkin, S. Yin, and L. Chen, "Optimization of nitrogen, phosphorus, and potassium fertilization rates for overseeded perennial ryegrass turf on dormant bermudagrass in a transitional climate," *Front. Plant Sci.*, vol. 9, p. 323819, Apr. 2018, doi: 10.3389/FPLS.2018.00487/BIBTEX.

[12] G. Sun, Y. Ding, X. Wang, W. Lu, Y. Sun, and H. Yu, "Nondestructive Determination of Nitrogen, Phosphorus and Potassium Contents in Greenhouse Tomato Plants Based on Multispectral Three-Dimensional Imaging," *Sensors 2019, Vol. 19, Page 5295*, vol. 19, no. 23, p. 5295, Dec. 2019, doi: 10.3390/S19235295.

[13] M. Safa, K. E. Martin, B. KC, R. Khadka, and T. M. R. Maxwell, "Modelling nitrogen content of pasture herbage using thermal images and artificial neural networks," *Therm. Sci. Eng. Prog.*, vol. 11, pp. 283–288, Jun. 2019, doi: 10.1016/J.TSEP.2019.04.005.

[14] Y. Peng *et al.*, "Estimation of leaf nutrition status in degraded vegetation based on field survey and hyperspectral data," *Sci. Reports 2020 101*, vol. 10, no. 1, pp. 1–12, Mar. 2020, doi: 10.1038/s41598-020-61294-7.

[15] Y. Peng, L. Zhao, Y. Hu, G. Wang, L. Wang, and Z. Liu, "Prediction of Soil Nutrient Contents Using Visible and Near-Infrared Reflectance Spectroscopy," *ISPRS Int. J. Geo-Information 2019, Vol. 8, Page 437*, vol. 8, no. 10, p. 437, Oct. 2019, doi: 10.3390/IJGI8100437.

[16] W. Frodella, G. Lazzeri, S. Moretti, J. Keizer, and F. G. A. Verheijen, "Applying Infrared Thermography to Soil Surface Temperature Monitoring: Case Study of a High-Resolution 48 h Survey in a Vineyard (Anadia, Portugal)," *Sensors 2020, Vol. 20, Page 2444*, vol. 20, no. 9, p. 2444, Apr. 2020, doi: 10.3390/S20092444.

[17] B. N. Aryalekshmi, R. C. Biradar, and J. Mohammed Ahamed, "Thermal imaging techniques in agricultural applications," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no.

Power System Technology
ISSN:1000-3673
Received: 06-01-2025 Revised: 15-02-2025 Accepted: 05-03-2025

12, pp. 2162–2168, Oct. 2019, doi: 10.35940/IJITEE.L2949.1081219.

[18] T. Hengl *et al.*, "Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning," *Nutr. Cycl. Agroecosystems*, vol. 109, no. 1, pp. 77–102, Sep. 2017, doi: 10.1007/S10705-017-9870-X/FIGURES/12.

[19] B. Dey, J. Ferdous, and R. Ahmed, "Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables," *Heliyon*, vol. 10, no. 3, p. e25112, Feb. 2024, doi: 10.1016/J.HELIYON.2024.E25112.

[20] M. Sujatha and C. D. Jaidhar, "Machine learning-based approaches to enhance the soil fertility—A review," *Expert Syst. Appl.*, vol. 240, p. 122557, Apr. 2024, doi: 10.1016/J.ESWA.2023.122557.

[21] F. Falcone, P. K. Mahapatro, R. Panigrahi, and N. Padhy, "Integrated Internet of Things and Artificial Intelligence System for Real-Time Multi-Nutrient Water Quality Analysis in Agriculture," *Eng. Proc. 2024, Vol. 82, Page 72*, vol. 82, no. 1, p. 72, Nov. 2024, doi: 10.3390/ECSA-11-20358.

[22] A. Sarangi, S. K. Raula, S. Ghoshal, S. Kumar, C. S. Kumar, and N. Padhy, "Enhancing Process Control in Agriculture: Leveraging Machine Learning for Soil Fertility Assessment," *Eng. Proc. 2024, Vol. 67, Page 31*, vol. 67, no. 1, p. 31, Sep. 2024, doi: 10.3390/ENGPROC2024067031.

[23] A. Munezero, A. Uwitonze, J. Kayalvizhi, C. Maniriho, J. Niyitegeka, and P. Ndorimana, "Machine Learning and Internet of Things Based Real Time NPK Fertilizer Prediction for Cassava Crop in Rwanda," *ACM Int. Conf. Proceeding Ser.*, pp. 12–17, Feb. 2024, doi: 10.1145/3651781.3651784.

[24] M. Awais *et al.*, "AI and machine learning for soil analysis: an assessment of sustainable agricultural practices," *Bioresour. Bioprocess. 2023 101*, vol. 10, no. 1, pp. 1–16, Dec. 2023, doi: 10.1186/S40643-023-00710-Y.

[25] M. K. Senapaty, A. Ray, and N. Padhy, "IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture," *Comput. 2023, Vol. 12, Page 61*, vol. 12, no. 3, p. 61, Mar. 2023, doi: 10.3390/COMPUTERS12030061.