RESEARCH ARTICLE Agricultural Science Digest, Volume 46 Special Issue: 107-118 (2026)

An Optimized Temporal Attention-based Deep Learning Framwork for Accurate Crop Yield Prediction from Multi-source Agricultural Data

M. Lokeshwari[1], Nivethitha Manavalagan[2], P. Amrutha[3], S. Harinee[4], N.R. Divyasree[4]

10.18805/ag.D-6558

ABSTRACT

Background: In precision agriculture, crop yield prediction is crucial for supporting effective management of crops, timely settlement of farmers crop insurances, food security planning and serve as decision tool for sustainable agricultural achievements. However, prediction of crop yield is difficult due to the complex interaction behavior among diverse conditions in weather, soil properties, multispectral remote sensing observations and the dynamics in crop growth. Despite existing AI techniques like machine learning and deep learning models outperformed in predicting crop yield, they are often not able to capture the temporal dependencies. Furthermore, determining optimal hyperparameters for model training remains a challenging task, significantly influencing prediction accuracy, model convergence and generalization performance.

Methods: To overcome this, the present study aimed to predict the crop yield using temporal attention (TA) based Artificial bee colony (ABC) optimized Bi - LSTM model using multi source agricultural data. The proposed model learns the sequential and temporal dependencies from both past and future time steps simultaneously, from the historical weather variables, soil properties, remote sensing derived normalized difference vegetation indices (NDVI) and historical yield data. The ABC algorithm was employed to automatically optimize the model hyperparameters with minimum run time.

Result: The model is trained using district wise rice yield records of Tamil Nadu, historical weather variables, Sentinel-2 derived NDVI and soil properties. The dataset collected for major rice growing districts from 1995 to 2025 (30 years). The model shows the superior performance over conventional methods like random forest, support vector regression, deep neural network, convolutional neural network, long short-term memory models with lesser root mean square error (RMSE = 298.01 kg ha^{-1}) and higher coefficient of determination, R^2 (81.76%). Also, the training time of TA - ABC - Bi-LSTM model is comparatively less than the comparative models because of ABC optimization algorithm tuned the best hyperparameters compared to traditional trial and error manual method.

Key words: Artificial bee colony, Bi-LSTM, Crop yield prediction, Temporal attention.

INTRODUCTION

Assessing crop yield is necessary to measure and improve the productivity of small and marginal farm holding through the proper decision making, effective crop management, market regulation and global food security planning (Xiao *et al.,* 2025). Reliable yield prediction is also required for taking decision on settling the crop insurance (Sadeh *et al.,* 2024). Early-season prediction of crop yield also assists researchers in optimizing irrigation, fertilizer application, pest and disease management and harvesting operations, thereby improving productivity and reducing production risks (Gavasso-Rita *et al.,* 2023). However, crop yield is dependent on numerous interacting factors, including weather variability, soil properties, crop genetics, agronomic practices, pest and disease incidence and extreme climatic events (Ansarifar *et al.,* 2021; Ishaq *et al.,* 2025; Jabed and Murad, 2024). The influence of these factors varies across regions, crop growth stages and growing seasons, resulting in substantial spatial and temporal variability in crop productivity (Ansarifar *et al.*, 2021; Nyéki and Neményi, 2022; Ishaq *et al.,* 2025).

Conventional crop yield predictions have several limitations in addressing this complexity. Empirical models shows statistical relationships between indices and yield but are often limited to specific crops, regions and growing seasons (Ishaq *et al.,* 2025; Khidirova and Karakus, 2026;

[1]Department of Agricultural Statistics, Division of Agricultural Economics, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore-641 114, Tamil Nadu, India. [2]Department of Seed Science and Technology, Division of Genetics and Plant Breeding, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore-641 114, Tamil Nadu, India. [3]Division of Plant Pathology, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore-641 114. Tamil Nadu, India. [4]Division of Horticulture, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore-641 114, Tamil Nadu, India.

Corresponding Author: M. Lokeshwari, Nivethitha Manavalagan; Department of Agricultural Statistics, Division of Agricultural Economics, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore-641 114, Tamil Nadu, India; Department of Seed Science and Technology, Division of Genetics and Plant Breeding, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore-641 114, Tamil Nadu, India.

Email: lokeshwarisumathi@gmail.com; nivethithamanas96@gmail.com

ORCIDs: https://orcid.org/0009-0000-1672-5311, https://orcid.org/0000-0003-4308-6847, https://orcid.org/0000-0003-4563-1669, https://orcid.org/0009-0004-5476-2446, https://orcid.org/0009-0004-4501-362X

How to cite this article: Lokeshwari, M., Manavalagan, N., Amrutha, P., Harinee, S. and Divyasree, N.R. (2026). An Optimized Temporal Attention-based Deep Learning Framework for Accurate Crop Yield Prediction from Multi-source Agricultural Data. *Agricultural Science Digest*. **46(Special Issue):** 107-118. doi: 10.18805/ag.D-6558.

Submitted: 22-07-2026 **Accepted:** 13-08-2026 **Online:** 10-09-2026

Nejadshamsi *et al.,* 2026). Process-oriented crop simulation models, such as DSSAT, APSIM, WOFOST and AquaCrop can stimulate crop growth under varying climatic conditions but require extensive data on weather, soil, crop management and cultivar characteristics as well as those are crop specific (Ashfaq *et al.*, 2025; Xiao *et al.,* 2025). Moreover, model calibration and validation across different regions and seasons are time-consuming and computationally demanding (Demissie *et al.,* 2026). These limitations restrict the scalability and operational use of existing approaches for large-scale agricultural applications. Consequently, there is an increasing demand for scalable, data-driven prediction frameworks capable of integrating heterogeneous agricultural datasets collected from multiple sources and across different crop growth stages (Duraisamy *et al.,* 2026).

Remote sensing and artificial intelligence (AI) has become an essential component of smart farming by providing timely information for data-driven agricultural decision-making (Phang *et al.,* 2023). Satellite platforms such as Sentinel-2A offer multispectral imagery for vegetation (Nevavuori *et al.,* 2019). The integration of satellite derived indices like Normalized Vegetative Indices (NDVI) has shown significant potential for estimating crop biophysical parameters (Näsi *et al.,* 2017) as well as their contribution to precision yield by integrating weather variables has emerged as an effective approach for improving crop yield prediction (Lokeshwari *et al.*, 2025; Lokeshwari *et al.,* 2026). Simultaneously, machine learning (ML) algorithms, including random forest (RF), Support Vector Regressor (SVR) and neural network-based models, have shown superior capability in modeling complex nonlinear relationships within agricultural datasets (Breiman, 2001; Cheng *et al.,* 2022). More recently, deep learning (DL) architectures such as deep neural networks (DNNs), convolutional neural networks (CNNs) and long short-term memory (LSTM) networks have further improved prediction accuracy by automatically learning hierarchical feature representations and temporal dependencies from high-dimensional data (LeCun *et al.,* 2015; Kalmani *et al.*, 2025; Gimba and Mishra, 2026). Nevertheless, their predictive performance remains highly dependent on optimal network architecture and hyperparameter selection, motivating the adoption of metaheuristic optimization techniques for efficient model tuning (Akay *et al.,* 2022).

Although there have been advancements in this field, there are still certain issues faced in the development of strong crop yield prediction models. Agricultural data is inherently heterogeneous in nature, consisting of observations based on remote sensing, climatic features, soil attributes and yield observations collected at various temporal and spatial scales. Deep learning techniques generally give equal weights to each observation in a time-series, which restricts their capability of determining the most impactful crop growth stages. In addition, hyperparameters tuning manually is computationally expensive and could result in sub-optimal performance of the model.

To solve all those key challenges, this paper presents a novel optimized temporal attention-based deep learning framework for accurate crop yield prediction from multi-source agricultural data. The developed framework combines veg

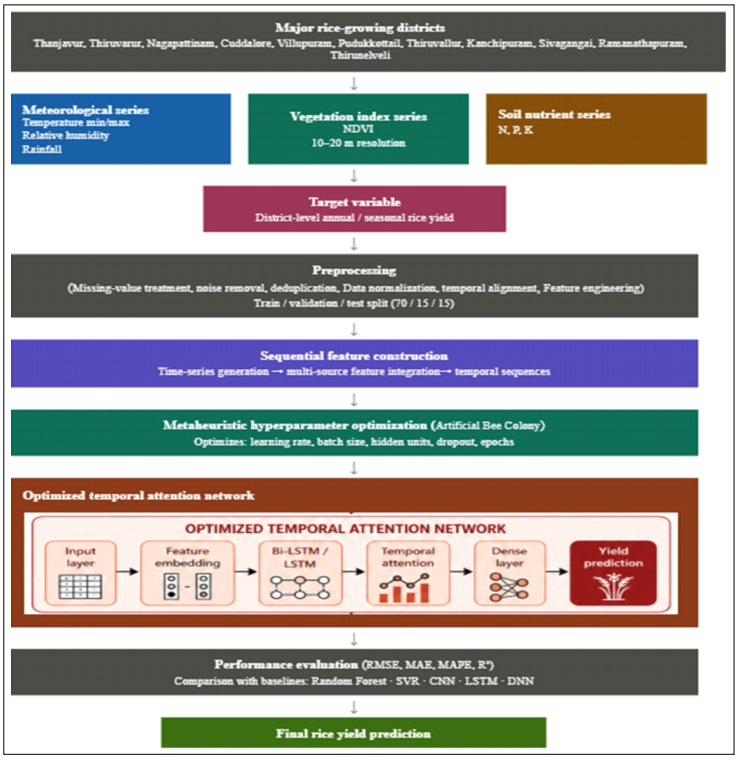


Fig 1: The overall architecture of proposed framework.

etation indices based on remote sensing, meteorological data, soil information and yield history into a Bi-directional Long Short-Term Memory (Bi-LSTM) neural network complemented by the temporal attention mechanism which helps to determine the most valuable periods of crops' development. Besides, the Artificial Bee Colony (ABC) algorithm is applied for automatic optimization of network parameters. The efficiency of the suggested framework will be verified through multi-source agricultural data testing against conventional machine learning and deep learning approaches in terms of performance metrics.

MATERIALS AND METHODS

The overall workflow of the proposed temporal attention-based artificial bee colony optimized bidirectional long short-term memory (TA-ABC-BiLSTM) framework is illustrated in Fig 1. The methodology comprises data collection, preprocessing, sequential feature construction, ABC-based hyperparameter optimization, Bi-LSTM with temporal attention and model evaluation.

Data collection

Our current study focuses on district level rice yield prediction in Tamil Nadu, India using agricultural datasets collected between 1995 and 2025 (Fig 2). Major rice growing districts were selected namely Thanjavur, Thiruvarur, Nagapattinam, Cuddalore, Villupuram, Pudukkottai, Tiruvallur, Kanchipuram, Sivagangai, Ramanathapuram and Tirunelveli. The details about the dataset and their source are mentioned in Table 1 also the summary statistics of target variable are given in Table 2.

Data preprocessing

The collected datasets were preprocessed to improve data quality and ensure consistency. Missing observations were imputed using the mean imputation method, followed by Min-Max normalization to standardize the variable scales. Weather variables, vegetation index, soil and historical yield datasets were then aligned according to crop growth stages to generate sequential input features. Following preprocessing, temporal alignment and sequence generation, the final

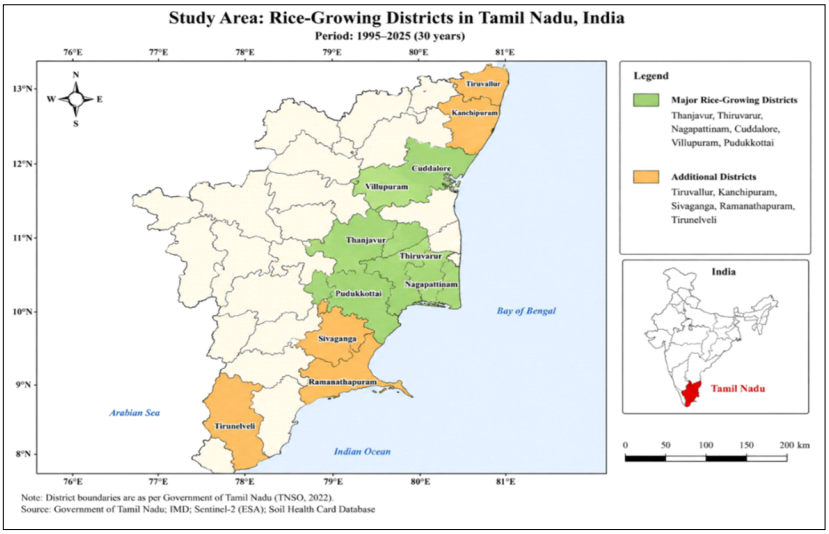


Fig 2: Study area showing the selected rice-growing districts in Tamil Nadu, India.

Table 1: Multi-source agricultural datasets used in the proposed framework.

Data type	Variables	Data source	Spatial resolution and temporal resolution
Meteorological data	Maximum temperature (Tmax), Minimum temperature (Tmin), Rainfall, Relative humidity (RH)	https://power.larc.nasa.gov/data-access-viewer/	District/Station level and Annual
Vegetation indices data	NDVI (Vegetation Indices)	Sentinel-2 MSI (ESA) processed using Google Earth Engine (GEE)	10-20 mb and Major growth stages
Soil data	Nitrogen (N), Phosphorus (P), Potassium (K),	Tamil Nadu Agricultural Department, Soil Health Card Database, NBSS and LUP	District/Field level and Seasonal
Crop yield data (Target variable)	District-wise annual/seasonal rice yield (kg ha^{-1})	Directorate of Economics and Statistics (DES), Government of Tamil Nadu	District level and Annual

dataset consisted of 1,023 multivariate temporal sequences, which were used for model development. The dataset was randomly divided into training (70%; n = 716), validation (15%; n = 153) and testing (15%; n = 154) subsets.

Sequential feature construction

To preserve the temporal dynamics of crop growth, the preprocessed data were transformed into multivariate temporal sequences prior to training. Weather variables (maximum temperature, minimum temperature, rainfall and relative humidity), Sentinel-2-derived NDVI, soil properties and historical rice yield were organized according to the major crop growth stages, namely Emergence, Vegetative, Flowering and Maturity. At each growth stage, observations from all data sources were combined to form a multivariate feature vector representing the crop condition during that period. The ordered feature vectors were subsequently arranged into a three-dimensional input tensor and used as input to the proposed TA-ABC-BiLSTM model (Fig 3).

The temporal input sequence is represented as.

$$X = (x_1, x_2, ..., x_T)$$

Where;

$$X_t = [x_t^{(1)}, x_t^{(2)}, ..., x_t^{(d)}]^T$$

and

$$X \in \mathbb{R}^{N \times T \times D}$$

Where,

N = Number of samples.

T = Number of crop growth stages (time steps).

d = Number of features at each stage.

Artificial bee colony (ABC)

The artificial bee colony (ABC) algorithm (Karaboga, 2005) was employed to optimize the hyperparameters of the proposed TA-Bi-LSTM model. ABC is a population-based metaheuristic optimization technique inspired by the foraging behavior of honeybee colonies, where employed bees, onlooker bees and scout bees collaboratively explore and exploit the search space to identify optimal solutions (Ghafil and Jármai, 2018; Srivastava *et al.*, 2021; Erkan *et al.,* 2023). Candidate solutions consisted of combinations of learning rate, batch size, hidden neurons, dropout rate and training epochs. Each solution was evaluated using the validation RMSE and the optimization process iteratively updated candidate solutions until convergence. The optimized hyperparameter values obtained through ABC were subsequently used to train the final prediction model. The search ranges of optimized hyperparameters are presented in Table 3.

Bi-LSTM

The bidirectional long short-term memory (Bi-LSTM) network was used to model temporal dependencies in historical weather variables, Sentinel-2-derived NDVI, soil properties and district-level rice yield. Unlike conventional LSTM, Bi-LSTM processes the input sequence in both forward and backward directions, enabling simultaneous utilization

Table 2: Summary statistics of the target variable.

Summary statistics	Paddy
Year range	1995-2025
Mean yield (kg ha^{-1})	3447.82
Standard deviation of yield	781.65
Minimum yield (kg ha^{-1})	324
Maximum yield (kg ha^{-1})	5472

Table 3: Hyperparameter search space used in the artificial bee colony optimization.

Hyperparameter	Search range
Learning rate	0.0001-0.01
Batch size	16-256
Hidden units	32-512
Dropout rate	0.0-0.5
Epochs	50-300
Optimizer	Adam/RMSprop

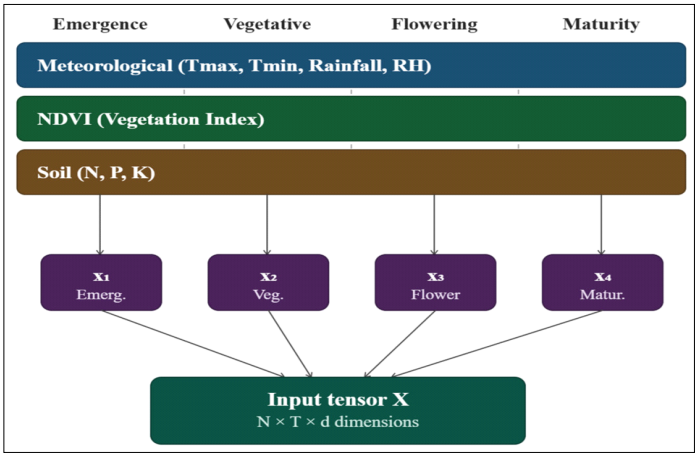


Fig 3: Construction of multivariate temporal sequences from multi-source agricultural data for input to the proposed deep learning model.

of past and future temporal information (Hochreiter and Schmidhuber, 1997; Schuster and Paliwal, 1997). The forward and backward hidden states were concatenated to generate a comprehensive temporal representation, which served as the input to the temporal attention layer.

The forward and backward hidden states are expressed as:

$$\overrightarrow{h_t} = LSTM_f(X_t, \overrightarrow{h_{t-1}})$$

$$\overleftarrow{h_t} = LSTM_b(x_t, \overleftarrow{h_{t-1}})$$

The final hidden representation is obtained by concatenating the two hidden states:

$$h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}]$$

Where,

$\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ = Forward and backward hidden states, respectively.

h_t = Combined temporal feature vector used as input to the temporal attention layer.

Attention mechanism

The attention mechanism models the dependencies between a collection of values and a target query by adaptively assigning an importance weight to each value. These weights are determined based on the similarity between the query and the corresponding keys, enabling the model to focus on the most relevant information for prediction.

The self-attention mechanism is a specialized form of the attention mechanism in which the query (Q), key (K) and value (V) are obtained through linear transformations of the same input sequence. This enables the model to effectively capture internal dependencies and contextual relationships within the data input. The self-attention mechanism is defined as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad ...(1)$$

Where,

$$Q = XW_Q, K = XW_K, V = XW_V$$

Where,

W_Q, W_K and W_V = Learnable weight matrices;

Q, K and V = Query, key and value matrices, respectively.

d_k = Dimension of the key vectors used to scale the dot-product attention scores.

Performance evaluation metrics

The performance of the proposed model was evaluated using mean absolute error (MAE), mean squared error (MSE), root Mean square error (RMSE) and mean absolute percentage error (MAPE).

Mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \quad ...(2)$$

Mean squared error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad ...(3)$$

Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad ...(4)$$

Mean absolute percentage error (MAPE):

$$MAPE = \frac{100}{N}\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad ...(5)$$

RESULTS AND DISCUSSION

The temporal variation in rice yield, weather variables (temperature, rainfall and relative humidity) and soil parameters (S1-S3) across the eleven rice-growing districts of Tamil Nadu is presented in Fig 4. Z-score standardization enabled comparison of variables measured on different scales by highlighting their temporal patterns. Rainfall exhibited the highest interannual variability, whereas soil parameters remained relatively stable throughout the study period. Crop yield generally followed fluctuations in rainfall and relative humidity, emphasizing the importance of moisture availability for rice production. In contrast, increasing maximum and minimum temperatures observed in several districts may adversely affect crop productivity by increasing evapotranspiration and shortening the grain-filling period. Districts such as Thanjavur, Thiruvarur and Nagapattinam showed strong synchronization between rainfall and yield, reflecting their dependence on monsoon rainfall, whereas Ramanathapuram and Sivagangai exhibited greater climatic variability because of their drought-prone conditions. The relatively stable soil parameters indicate that annual yield fluctuations were primarily influenced by climatic variability rather than abrupt changes in soil fertility.

The Pearson correlation matrix (Fig 5) further supported these observations. Rainfall and relative humidity showed positive correlations with rice yield, indicating their beneficial influence on crop growth, whereas maximum temperature exhibited a weak negative correlation, suggesting adverse effects under higher temperatures. Maximum and minimum temperatures were positively correlated owing to their seasonal patterns, while relative humidity was negatively associated with temperature. Soil parameters showed moderate correlations with yield, implying that their influence is largely mediated through climatic conditions rather than acting independently. These observations are consistent with previous studies that reported weather variables and remotely sensed vegetation indices as major determinants of crop yield variability (Ishaq *et al.,* 2025; Xiao *et al.,* 2025).

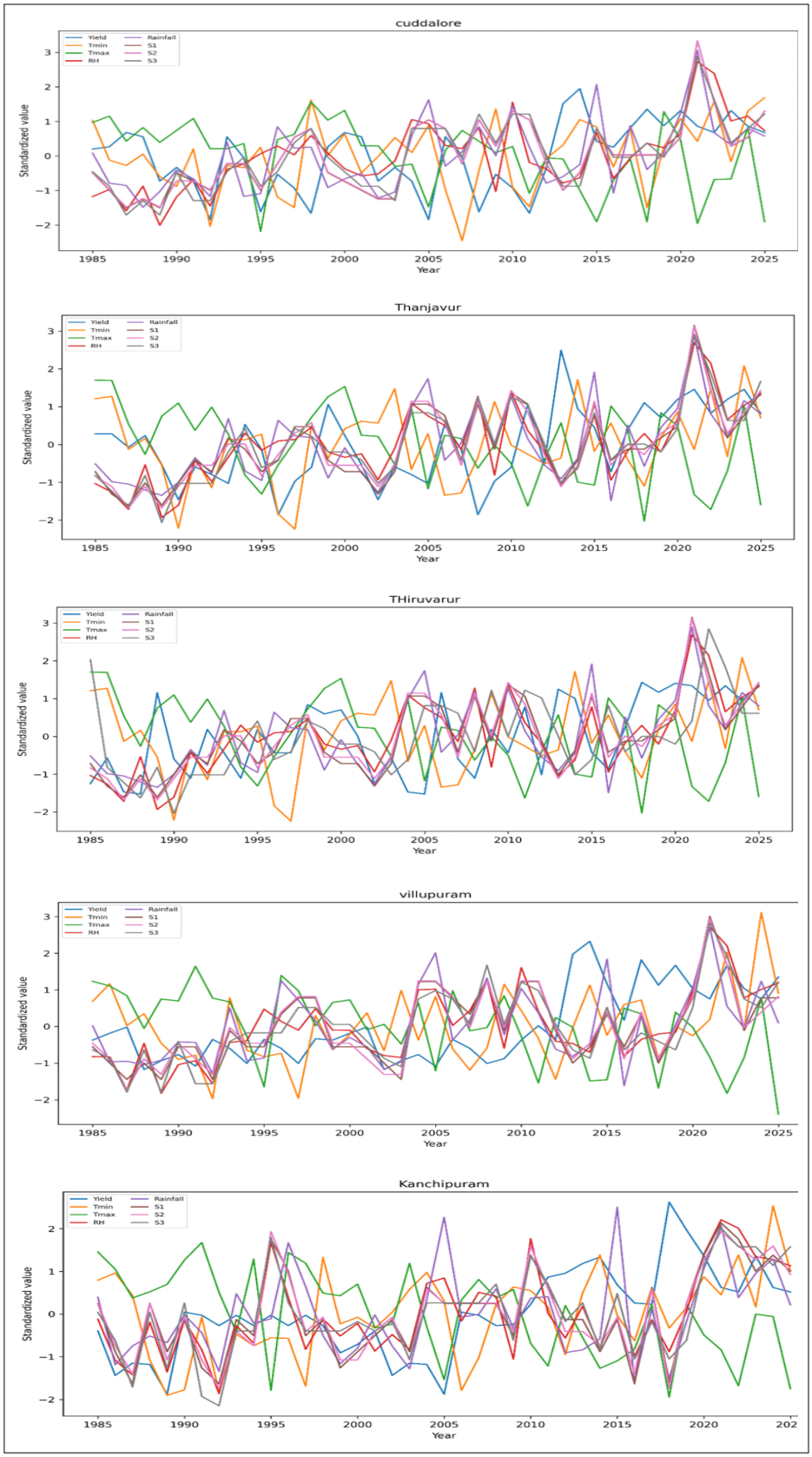


Fig 4: Continue...

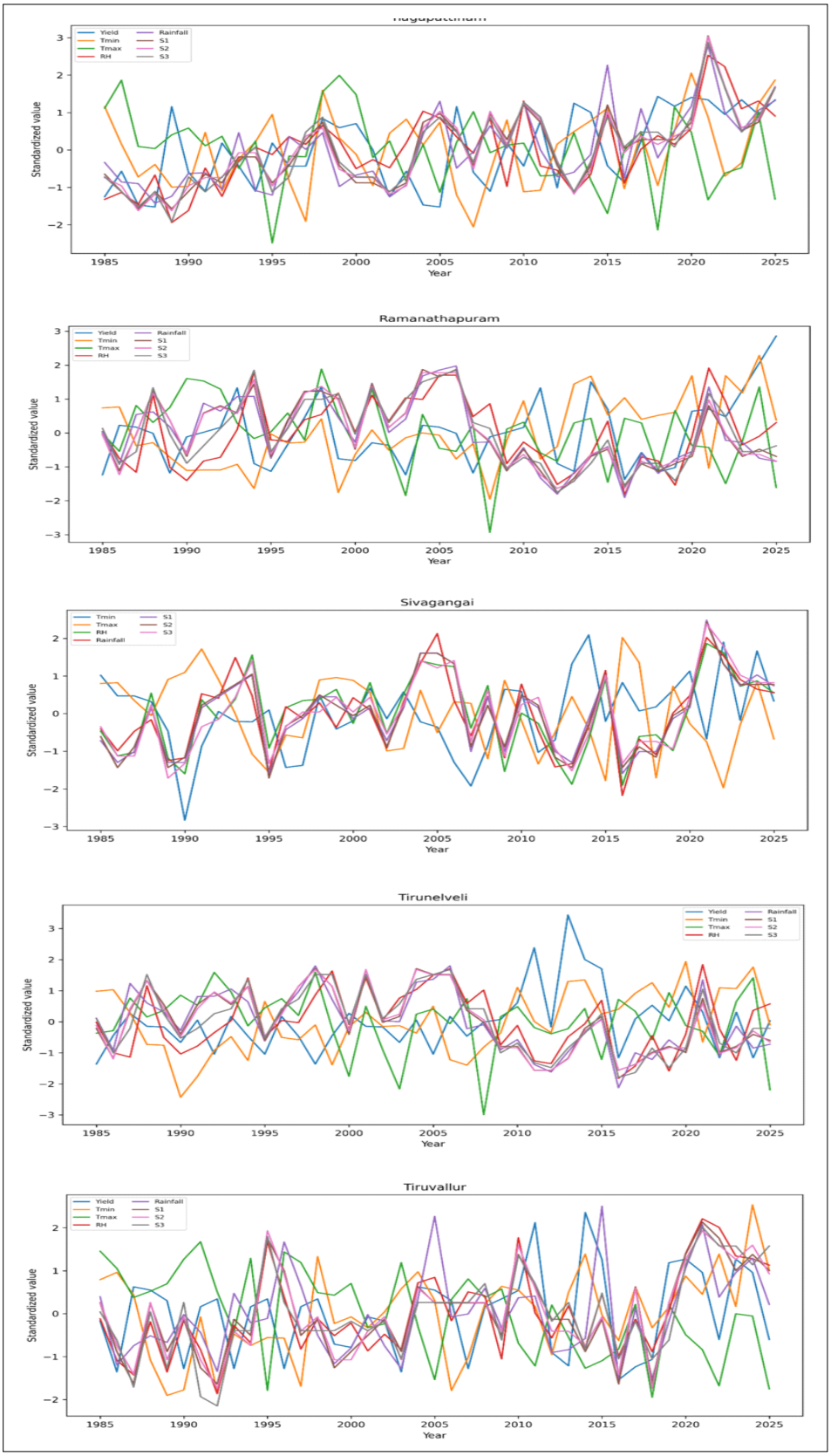


Fig 4: Continue...

Fig 4: Continue...

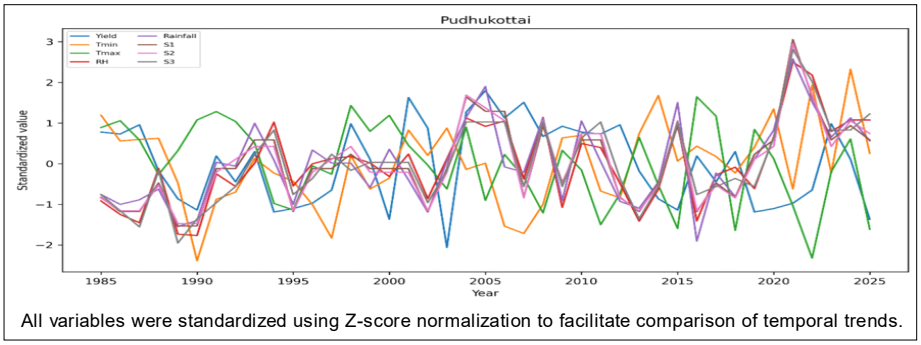


Fig 4: District-wise temporal variations of crop yield, minimum temperature (Tmin), maximum temperature (Tmax), relative humidity (RH), rainfall and soil parameters (S1-S3) during the study period.

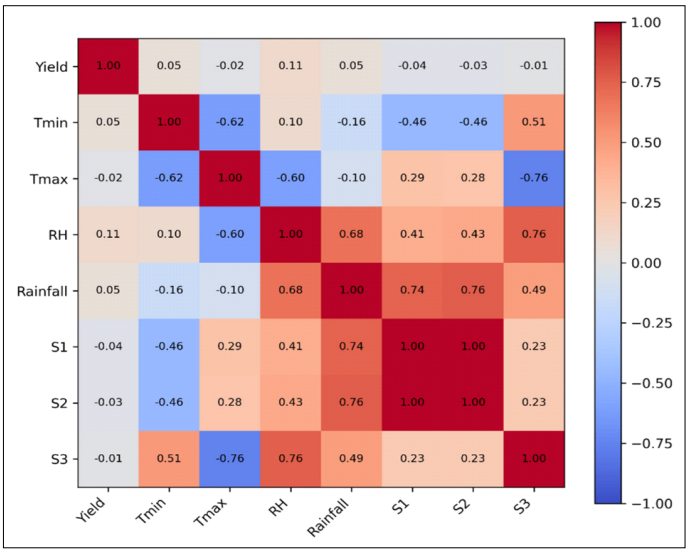


Fig 5: Overall Pearson correlation heatmap illustrating the relationships among crop yield, climatic variables and soil parameters across the eleven study districts.

The proposed TA-ABC-Bi-LSTM model was implemented in Python using the TensorFlow framework. Random Forest (RF), Support Vector Regressor (SVR), Deep Neural Network (DNN), Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models were developed as benchmark methods using identical training and testing datasets. To ensure a fair comparison, the hype rparameters of all models were optimized using the Artificial Bee Colony (ABC) algorithm. The implementation parameters of the proposed and baseline models are given in Table 4. The optimized hyperparameter configuration of the proposed model is summarized in Table 5. As illustrated in Fig 6, the validation RMSE decreased steadily during successive optimization cycles before converging to a stable minimum, demonstrating the ability of the ABC algorithm to efficiently explore the search space and identify an optimal hyperparameter combination while avoiding premature convergence.

Table 4: Implementation parameters of the proposed and baseline models.

Parameter	Value/Setting
Optimizer (network weights)	Adam
Hyperparameter optimizer	Artificial bee colony (ABC)
ABC population size (Food sources)	30
ABC max cycles	100
ABC limit (abandonment)	20
Loss function	Mean squared error (MSE)
Early stopping patience	15 epochs
Train / Validation / Test Split	70% / 15% / 15%
Sequence length (T)	6 (growth stages)
Feature dimension (d)	10
Random seed	42

The temporal attention mechanism assigned different importance to individual crop growth stages (Fig 7). The flowering and grain-filling stages received the highest attention weights, indicating that these reproductive stages contributed most to yield prediction. This observation agrees with rice physiology, where assimilating partitioning

and grain development during reproductive growth largely determines the final grain yield. Furthermore, recent attention-based deep learning studies have demonstrated that assigning adaptive weights to critical phenological stages substantially improves the learning of temporal dependencies and enhances prediction accuracy compared with conventional recurrent neural networks (Jia *et al.*, 2024; Mirhoseini Nejad *et al.,* 2024).

The learning behavior of the proposed model is illustrated in Fig 8 and 9. Training and validation loss, together

Table 5: Optimal hyperparameters selected using the artificial bee colony algorithm.

Hyperparameter	Initial value	Optimized value
Learning rate	0.001	0.0006
Batch size	32	64
Hidden units	64	128
Dropout	0.20	0.30
Epochs	100	150
Optimizer	Adam	Adam

Table 6: Comparative performance of machine learning and deep learning models.

Model	RMSE	MAE	MAPE	R^2
Random Forest	512.40	395.10	16.84	61.20
SVR	478.90	368.45	15.72	65.80
DNN	421.30	321.84	13.65	70.40
CNN	389.60	297.26	12.48	74.10
LSTM	342.70	258.43	10.76	78.30
Proposed model	298.01	223.75	9.18	81.76

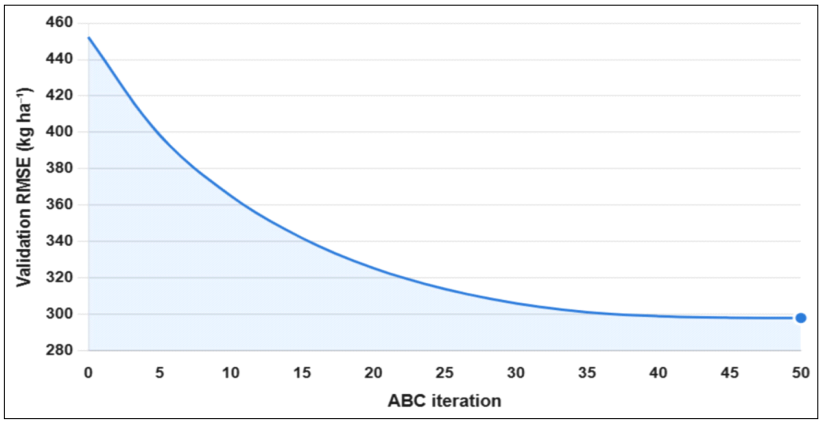


Fig 6: Convergence behavior of the artificial bee colony algorithm.

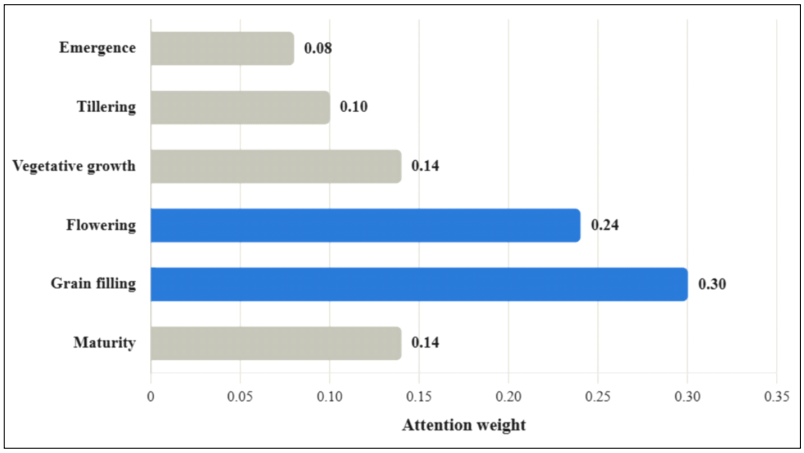


Fig 7: Attention weights across the six crop growth stages, with flowering and grain filling highlighted as the stages the model weights most heavily.

with RMSE, decreased rapidly during the initial epochs and gradually stabilized after approximately 60-70 epochs, indicating efficient model convergence. The close agreement between the training and validation curves suggests good generalization with no evidence of overfitting. Furthermore, early stopping effectively prevented unnecessary training iterations, improving computational efficiency while maintaining prediction performance.

The predictive performance of the proposed framework is compared with benchmark models in Table 6. The TA-ABC-Bi-LSTM model achieved the lowest RMSE (298.01 kg ha^{-1}), demonstrating superior prediction accuracy over RF, SVR, DNN, CNN and conventional LSTM models. The comparatively poor performance of RF can be attributed to its inability to explicitly model temporal dependencies, whereas SVR struggled to represent the complex nonlinear

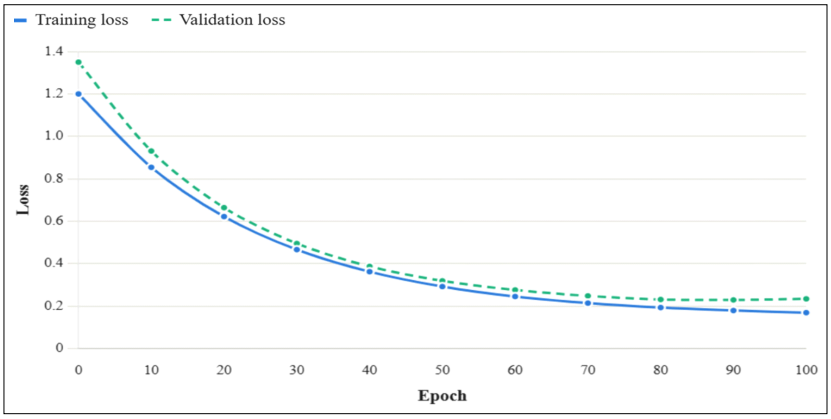


Fig 8: Training and validation loss curves.

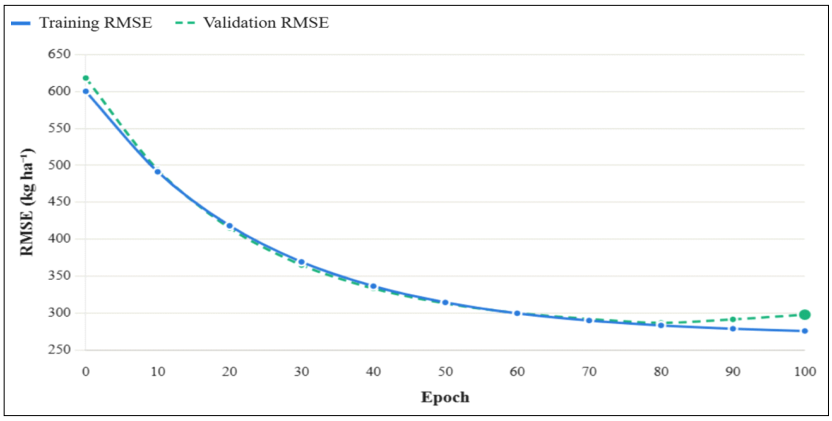


Fig 9: Training and validation RMSE curves.

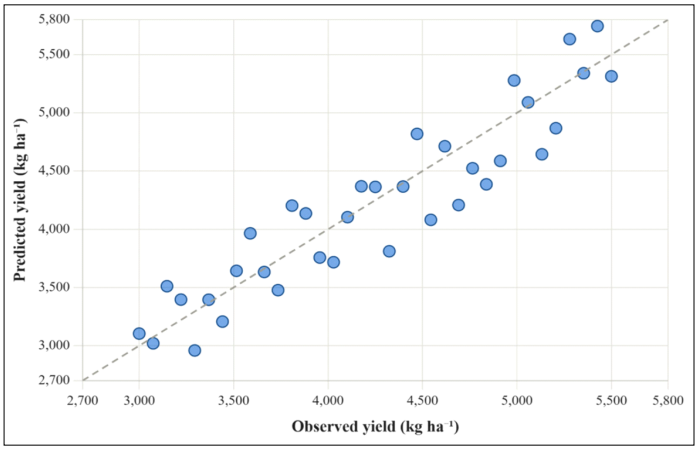


Fig 10: Observed vs. Predicted rice yield, with the 1:1 reference line showing perfect prediction.

interactions among weather, soil, vegetation and temporal variables. Although DNN and CNN captured nonlinear relationships and local feature patterns, they were less effective in learning long-term temporal dependencies. Conventional LSTM improved prediction accuracy through its memory architecture but assigned similar importance to all historical observations. In contrast, the proposed framework integrates bidirectional temporal learning with a temporal attention mechanism that automatically identifies the most influential crop growth stages, while the ABC algorithm efficiently determines the optimal hyperparameter combination. These complementary components collectively contributed to improved prediction accuracy, faster convergence and better model generalization. Furthermore, the obtained results are consistent with recent studies demonstrating the advantages of attention-based deep learning and metaheuristic optimization for crop yield prediction using multi-source agricultural datasets (Lokeshwari *et al.,* 2026; Demissie *et al.,* 2026).

The observed vs. predicted rice yields is presented on Fig 10. Most observations were distributed close to the 1:1 reference line, indicating high predictive accuracy across districts and growing seasons. The limited deviation observed under a few extreme climatic conditions demonstrates the robustness of the proposed framework in capturing spatial and temporal variability in rice yield under diverse agro-climatic environments.

CONCLUSION

The present study proposed a novel Temporal Attention-Based Artificial Bee Colony Optimized Bidirectional Long Short-Term Memory (TA-ABC-Bi-LSTM) framework for district-level crop yield prediction using multi-source agricultural data. By integrating a temporal attention mechanism with the Bi-LSTM architecture and employing the Artificial Bee Colony (ABC) algorithm for automatic hyperparameter optimization, the proposed framework effectively captured temporal dependencies and identified the most influential crop growth stages contributing to crop yield prediction. The experimental results demonstrated that the proposed model outperformed conventional machine learning and deep learning models, achieving superior predictive accuracy with lower prediction error and improved model generalization. Furthermore, the integration of meteorological variables, Sentinel-2-derived NDVI, soil properties and historical yield data into a unified temporal representation enhanced the model's capability to learn the complex nonlinear relationships influencing crop productivity.

The proposed framework provides a reliable decision-support tool for precision agriculture with potential applications in early crop yield forecasting, resource allocation, agricultural planning and crop insurance assessment. Nevertheless, the present study was conducted using district-level historical datasets with a hold-out validation strategy, which may limit its applicability to field-level prediction and other agro-climatic regions. Future research will focus on validating the proposed framework using independent datasets and cross-validation techniques, incorporating higher-resolution remote sensing data, UAV and hyperspectral imagery, Transformer-based deep learning architectures and Explainable Artificial Intelligence (XAI) methods to further improve the robustness, interpretability and generalizability of crop yield prediction models.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this article. No funding or sponsorship influenced the design of the study, data collection, analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Akay, B., Karaboga, D. and Akay, R. (2022). A comprehensive survey on optimizing deep learning models by metaheuristics. *Artificial Intelligence Review*. **55:** 829-894. https://doi.org/10.1007/s10462-021-09992-0.

Ansarifar, J., Wang, L. and Archontoulis, S.V. (2021). An interaction regression model for crop yield prediction. *Scientific Reports*. **11:** 17754. https://doi.org/10.1038/s41598-021-97221-7.

Ashfaq, M., Khan, I., Afzal, R.F., Shah, D., Ali, S. and Tahir, M. (2025). Enhanced wheat yield prediction through integrated climate and satellite data using advanced AI techniques. *Scientific Reports*. **15:** 18093. https://doi.org/10.1038/s41598-025-02700-w.

Breiman, L. (2001). Random forests. *Machine Learning*. **45:** 5-32. https://doi.org/10.1023/A:1010933404324.

Cheng, E., Zhang, B., Peng, D., Zhong, L., Yu, L., Liu, Y., Xiao, C. *et al.* (2022). Wheat yield estimation using remote sensing data based on machine learning approaches. *Frontiers in Plant Science*. **13:** 1090970. https://doi.org/10.3389/fpls.2022.1090970.

Demissie, W.A., Sebastiani, L. and Rossetto, R. (2026). Integration of artificial intelligence and remote sensing for crop yield prediction and crop growth parameter estimation in Mediterranean agroecosystems: Methodologies, emerging technologies, research gaps and future directions. *European Journal of Agronomy*. **173:** 127894. https://doi.org/10.1016/j.eja.2025.127894.

Duraisamy, P., Sriharani, R., Harish, D.V. and Jones Martin, J. (2026). A Multi-source Data-driven Hybrid Model for Predicting Crop Yield under Varying Climatic Conditions. Proceedings of 5th International Conference on Communication, Computing and Electronics Systems, ICCCES 2026. 1055-1060. https://doi.org/10.1109/ICCCES62661.2026.11436480.

Erkan, U., Toktas, A. and Ustun, D. (2023). Hyperparameter optimization of deep CNN classifier for plant species identification using artificial bee colony algorithm. *Journal of Ambient Intelligence and Humanized Computing*. **14:** 8827-8838. https://doi.org/10.1007/s12652-021-03631-w.

Gavasso-Rita, Y.L., Papalexiou, S.M., Li, Y., Elshorbagy, A., Li, Z. and Schuster-Wallace, C. (2023). Crop models and their use in assessing crop production and food security: A review. *Food and Energy Security*. **13(1):** e503. https://doi.org/10.1002/fes3.503.

Ghafil, H.N. and Jármai, K. (2018). Comparative Study of Particle Swarm Optimization and Artificial Bee Colony Algorithms. MultiScience - XXXII. microCAD International Multidisciplinary Scientific Conference. https://doi.org/10.26649/musci.2018.030

Gimba, M. A. and Mishra, K.P (2026). Explainable AI-enabled mixed-crop recommendation model using hybrid machine learning and ant colony optimization. *Indian Journal of Agricultural Research.* **60(6):** 936-942. doi: 10.18805/IJARe.A-6497.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*. **9(8):** 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Ishaq, R.A.F., Zhou, G., Jing, G., Shah, S.R.A., Ali, A., Imran, M., Jiang, H. and Obaid-Ur-Rehman. (2025). Geospatial robust wheat yields prediction using machine learning and integrated crop growth model and time-series satellite data. *Remote Sensing*. **17(7):** 1140. https://doi.org/10.3390/rs17071140.

Jabed, M.A. and Murad, M.A.A. (2024). Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. *Heliyon*. **10(24):** e40836. https://doi.org/10.1016/j.heliyon.2024.e40836.

Jia, Z., Wu, K., Wang, H., Zeng, W., Guo, Y. and Liang, D. (2024). An improved crop yield prediction using CNN-BiLSTM model with attention mechanism. *Journal of the ASABE*. **67(6):** 1459-1467. https://doi.org/10.13031/ja.15629.

Kalmani, H.V, Dharwadkar, V.N and Thapa, V. (2025). Crop yield prediction using deep learning algorithm based on CNN-LSTM with attention layer and skip connection. *Indian Journal of Agricultural Research.* **59(8):** 1303-1311. doi: 10.18805/IJARe.A-6300.

Karaboga, D. (2005). An Idea Based on Honeybee Swarms for Numerical Optimization. Technical Report TR06, Erciyes University, Engineering Faculty, Computer Engineering Department, 2005.

Khidirova, E. and Karakus, O. (2026). UniCrop: A universal, multi-source data engineering pipeline for scalable crop yield prediction. *Applied Sciences.* **16(10):** 4724. https://doi.org/10.3390/app16104724.

LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. *Nature*. **521:** 436-444. https://doi.org/10.1038/NATURE14539.

Lokeshwari, M., Jha, G.K., Lama, A., Praveenkumar, A., Kumari, J., Yashavantha Kumar, K.J. and Parsad, R. (2025). Wheat yield prediction through artificial bee colony-enhanced convolutional neural network. *Indian Journal of Genetics and Plant Breeding*. **85(1):** 78-86. https://doi.org/10.31742/ISGPB.85.1.9.

Lokeshwari, M., Jha, G.K., Praveenkumar, A., Kumari, J., Navathe, S., Singh, G.P. and Prasad, P.V.V. (2026). A novel deep learning framework for field-scale wheat yield prediction. *Theoretical and Applied Genetics*. **139:** 79. https://doi.org/10.1007/s00122-026-05166-0.

Mirhoseini Nejad, S.M., Abbasi-Moghadam, D., Sharifi, A. and Tariq, A. (2024). Capsular attention Conv-LSTM network (CACN): A deep learning structure for crop yield estimation based on multispectral imagery. *European Journal of Agronomy*. **161:** 127369. https://doi.org/10.1016/j.eja.2024.127369.

Näsi, R., Viljanen, N., Kaivosoja, J., Hakala, T., Pandžic, M., Markelin, L. and Honkavaara, E. (2017). Assessment of various remote sensing technologies in biomass and nitrogen content estimation using an agricultural test field. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. **42(3W3):** 137-141. https://doi.org/10.5194/isprs-archives-XLII-3-W3-137-2017.

Nejadshamsi, S., Zhang, Y., Porth, B., Zaki, S., Porth, L. and Khoshdel, V. (2026). FARM: Crop yield prediction *via* regression on Prithvi's encoder for satellite sensing. *AgriEngineering*. **8(1):** 2. https://doi.org/10.3390/agriengineering8010002.

Nevavuori, P., Narra, N. and Lipping, T. (2019). Crops yield prediction with deep convolutional neural networks. *Computers and Electronics in Agriculture*. **163:** 104859. 10.1016/j.compag.2019.104859.

Nyéki, A. and Neményi, M. (2022). Crop yield prediction in precision agriculture. *Agronomy*. **12(10):** 2460. https://doi.org/10.3390/agronomy12102460.

Phang, S.K., Chiang, T.H.A., Happonen, A. and Chang, M.M.L. (2023). From satellite to UAV-based remote sensing: A review on precision agriculture. *IEEE Access*. **11:** 127057-127076. 10.1109/ACCESS.2023.3330886.

Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J.P., Chen, Y. and Chenu, K. (2024). Versatile crops yield estimators. *Agronomy for Sustainable Development*. **44:** 42. https://doi.org/10.1007/s13593-024-00974-4.

Schuster, M. and Paliwal, K.K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing.* **45(11):** 2673-2681. https://doi.org/10.1109/78.650093.

Srivastava, A.K., Kumar, Y. and Singh, P.K. (2021). Artificial bee colony and deep neural network-based diagnostic model for improving the prediction accuracy of diabetes. *International Journal of E-Health and Medical Communications*. **12:** 32-50. https://doi.org/10.4018/IJEHMC.2021030102.

Xiao, G., Huang, J., Zhuo, W., Huang, H., Song, J., Du, K., Wang, J. *et al.* (2025). Progress and perspectives of crop yield forecasting with remote sensing: A review. *IEEE Geoscience and Remote Sensing Magazine*. **13(3):** 338-368. https://doi.org/10.1109/MGRS.2025.3571906.