



# Fixed-adaptive Temporal Attention Network for Predicting Crop Yield

S. Padmanayaki<sup>1</sup>, K. Geetha<sup>1</sup>

10.18805/ag.D-6468

## ABSTRACT

**Background:** In agriculture, Crop Yield Prediction (CYP) by Deep Learning (DL) models such as Deep Neural Network (DNN) and Convolutional Neural Network (CNN) is an emerging technique that utilizes weather, soil and previous year yield information for different crops. On the other hand, these models often struggle to effectively handle time-series data, which can lead to challenges in capturing temporal relationships between weather, soil and crop yield factors. This can affect the accuracy of predicting crop yields.

**Methods:** To address this issue, the Fixed-Adaptive Temporal Attention Network (FAT-AttNet) is proposed in this study for CYP using a time-series crop yield dataset, which includes historical information on soil, weather and crop yields. A normalization technique is used to preprocess the dataset, removing irregularities and empty values. Next, the FAT-AttNet model is trained to predict crop yields. This model consists of fixed and adaptive temporal modules. In the fixed temporal module, 18 parallel one-dimensional (1D) CNNs are used to process the stable factors, such as soil data, to capture temporal features. In the adaptive temporal module, 16 parallel two-dimensional (2D) CNNs capture the dynamic relationships among weather, soil and crop growth factors over time. The results from these models are combined using the lateral connection and an attention strategy is applied to adjust the weight of each factor for more accurate predictions. Furthermore, a Fully Connected (FC) layer and an output layer are used to predict crop yields.

**Result:** The experiments are conducted using an agricultural crop yield in Indian States dataset (ACY-ISD) to evaluate the performance of the FAT-AttNet model for five major crops in Tamil Nadu using the historical data from 2002 to 2023. The results show that the FAT-AttNet model achieves a Mean Absolute Error (MAE) of 0.05, 0.059, 0.061, 0.058 and 0.074 for groundnut, maize, moong, rice and urad yield predictions, respectively. The Root Mean Square Error (RMSE) for the groundnut, maize, moong, rice and urad yield predictions is 0.217, 0.268, 0.245, 0.235 and 0.266, respectively. The correlation coefficients for groundnut, maize, moong, rice and urad yield predictions are 0.862, 0.846, 0.864, 0.860 and 0.859, respectively. The results of FAT-AttCNN models outperform the Improved Optimizer Function with Long Short-Term Memory (IOF-LSTM), CNN-Bidirectional LSTM (BiLSTM), MayFly Algorithm empowered attention-BiLSTM (MFA-BiLSTM) and Spatio-Temporal Attention-based CNN (STACNN) models.

**Key words:** Attention strategy, CNN, Crop yield prediction, Fixed-adaptive temporal network, Temporal relationship.

## INTRODUCTION

Agriculture represents a significant social concern, serving as the fundamental source of food supply. Numerous countries continue to experience food shortages as a result of rapid population growth (Xia *et al.*, 2022). Crop yield is a widely used metric that signifies the quantity of agricultural output gathered per unit of land area. In conventional agriculture, farmers depend on their expertise and fundamental environmental variables, such as precipitation, temperature, soil composition and pesticide use, to forecast crop yields (Thimmegowda *et al.*, 2024). Leveraging historical data about these characteristics facilitates early crop forecasting, assisting farmers and agronomists in making educated harvesting choices. Gathering data and evaluating it externally is essential, which might be difficult for farmers who may not fully grasp the fundamental principles. Therefore, many researchers have developed various techniques, including stochastic modeling and Artificial Intelligence (AI), to accurately estimate yield quality and quantity for different crops (Zvobgo *et al.*, 2023; Al-Adhaileh and Aldhyani, 2022).

Contemporary CYP methods can be categorized into linear models, Machine Learning (ML) models and crop

<sup>1</sup>Department of Computer Science, Bharathiar University, Coimbatore-641 046, Tamil Nadu, India.

**Corresponding Author:** S. Padmanayaki, Department of Computer Science, Bharathiar University, Coimbatore-641 046, Tamil Nadu, India. Email: padmanayakiphd123@gmail.com

**How to cite this article:** Padmanayaki, S. and Geetha, K. (2026). Fixed-adaptive Temporal Attention Network for Predicting Crop Yield. *Agricultural Science Digest*. 1-9. doi: 10.18805/ag.D-6468.

**Submitted:** 03-11-2025 **Accepted:** 06-01-2026 **Online:** 26-01-2026

models. Linear models are understandable by measuring the additive weight of all factors. However, their prediction accuracy was low because they could not learn inherently nonlinear relationships among each factor. Crop models are a type of nonlinear model developed to predict crop yield (Luo *et al.*, 2023; Dwivedi *et al.*, 2022). These models offer explicit correlations between yield factors and weather conditions in multiple phases of the crop growth cycle. However, gathering yield data and adjusting model parameters can be time-consuming and labor-intensive. Additionally, prediction accuracy was low (Lin *et al.*, 2023).

Many researchers have utilized DNN and CNN models to forecast different crop yields using factors like soil, weather and yield information; some have validated results using traditional ML models (Rashid *et al.*, 2021). These developments depict the shift toward modern computational approaches to CYP. ML and DL models can reproduce intricate, nonlinear associations among numerous influencing variables, thus providing higher accuracy and strength of prediction than conventional techniques.

ML algorithms have been implemented for CYP, including random forest and decision trees (Shaikh *et al.*, 2022). However, the prediction accuracy of these algorithms was low for large datasets (Trentin *et al.*, 2024). These algorithms require different and independent algorithms for feature extraction and prediction tasks, leading to high computation time. Also, missing values in the dataset can impact the training of these algorithms and lead to unreliable predictions. To solve these problems, DL models have been employed for CYP in recent years (Talaie *et al.*, 2023). DL models can unify feature learning and prediction tasks in a single framework. They can model more complex relationships between yield and other factors, thus enhancing the discriminability of various features (Paudel *et al.*, 2023).

Recent studies have explored a variety of ML and DL approaches for CYP. A hybrid technique (Batool *et al.*, 2022) based on the AquaCrop simulation model and regression algorithms was developed to predict tea crop yield from weather, crop and soil factors, but MAE and RMSE remained high since these algorithms cannot capture the spatio-temporal relations among several factors. Seireg *et al.* (2022) applied an ensemble ML algorithm combining stacking and cascading regression to wild blueberry yield prediction using weather factors, yet RMSE and MAE were very high due to large datasets. Extra Tree and AdaBoost algorithms (Khan *et al.*, 2022) were employed for oil palm yield prediction using multisource data but suffered from low values. Srivastava *et al.*, (2022) used CNN models for winter wheat yield forecasting and Oikonomidis *et al.*, (2022) used hybrid CNN approaches for soybean yield recognition. However, they struggled with low correlation coefficients. Similarly, Functional Artificial Neural Network (FLANN) (Jena *et al.*, 2023), IOF-LSTM (Bhimavarapu *et al.*, 2023), CNN-BiLSTM (Saini *et al.*, 2023), MFA-BiLSTM (Krishna *et al.*, 2024) and STACNN-BiLSTM hybrids (Saravanan and Bhagavathiappan, 2024), CNN-LSTM with attention layer and skip connection (Kalmani *et al.*, 2025) faced relatively higher RMSE, MAE, or low correlation values, highlighting ongoing challenges in capturing complex and long-term dependencies in crop yield data.

### Problem definition

Even though DNN and CNN models are effective for CYP, they have shortcomings when applied to time-series data in this context. These models are not well-suited for handling crop yield data over time because they cannot

capture significant temporal correlations between various factors over a period. This is mostly significant in understanding long-term environmental (*i.e.*, weather) patterns to predict crop yields.

### Main contributions of the paper

In this paper, a new FAT-AttNet is developed for CYP using a time-series crop yield dataset, which includes historical information on soil, weather and crop yields. A normalization technique is used to preprocess the dataset, removing irregularities and empty values. Next, the FAT-AttNet model is trained to predict crop yields. This model consists of fixed and adaptive temporal modules. In the fixed temporal module, 18 parallel 1D CNNs are used to process the soil data and capture temporal features. In the adaptive temporal module, 16 parallel 2D CNNs are employed to capture the dynamic relationships among weather, soil and crop growth factors over time. Moreover, the results from these models are combined using the lateral connection and an attention strategy is applied to adjust the weight of each factor for more accurate predictions. Finally, the FC layer followed by an output layer is used for CYP. So, the FAT-AttNet approach successfully seizes fixed and adaptive temporal correlations among various factors to boost prediction accuracy.

## MATERIALS AND METHODS

This part presents a concise overview of the FAT-AttNet framework for CYP. The configuration of this research is depicted in Fig 1. At first, an agricultural crop yield dataset is acquired, which comprises yield data, weather data and soil data for various crops. Then, data preprocessing is applied to remove empty values and other outliers. Then, these data are given to the FAT-AttNet model to estimate crop yield. The estimated values are then used to evaluate the efficiency of FAT-AttNet. The following subsection describes the FAT-AttNet model briefly.

### Data collection

This research was conducted at the Department of Computer Science, Bharathiar University, Coimbatore, Tamil Nadu, India, during the years 2024 and 2025. The crop yield prediction experiment was implemented in MATLAB 2019b on a system with an Intel i5 processor and 4GB RAM. The dataset for this study was collected from the ACY-ISD. It contains information on various crops from different Indian states between 1997 and 2020. It is available at <https://www.kaggle.com/datasets/akshatgupta7/crop-yield-in-indian-states-dataset/data>. It includes essential attributes for CYP, such as crop types, years, harvesting periods, states, regions for farming, production measures, weather, soil quality (*e.g.*, fertilizer use, pesticide use) and intended yields. Among many crop types, this study focuses on predicting yield values of five major crops in Tamil Nadu, a South Indian state, including groundnut, maize, moong, rice and urad.

The dataset is split in an 80:20 ratio for training and testing. The dataset under study covers 1997–2020, but when it is limited to Tamil Nadu and the five chosen crops, the useful time series used in the present research is 2002-2023. A 22-year sequence is split into 80:20 train-test to indicate that the latest decade should be used (2015-2024) to test and visualize performances.

### Data pre-processing

This stage involves a normalization technique based on min-max scaling to convert the data to a specific range between 0 and 1. This can handle missing values and ensure uniformity across each feature (e.g., crop yield, soil and weather data). It is defined in Eq. (1).

$$\hat{x}_t^i = \frac{x_t^i - x_{\min}^i}{x_{\max}^i - x_{\min}^i} \quad \dots(1)$$

In Eq. (1),  $\hat{x}_t^i$  is the  $i^{\text{th}}$  feature at time  $t$ ,  $x_{\max}^i$  and  $x_{\min}^i$  are the maximum and minimum values of the corresponding feature. By using this process, the pre-processed crop yield dataset is obtained and used to train the FAT-AttNet model for CYP.

### Crop yield prediction using FAT-AttNet model

The aim of this study is CYP over a defined period utilizing past data. So, the input for the FAT-AttNet consists of a time series of length  $N$ , depicted by  $x_{t-N}, \dots, x_t$ .  $x_t$  in time-series represents a vector comprising observed values of yield, soil and weather data at time  $t$ . The FAT-AttNet model has a fixed temporal module and an adaptive temporal module. The fixed temporal unit is utilized to capture the temporal characteristics of the soil data and the dynamic temporal module is utilized to capture the dynamic relationships among weather, soil and crop growth factors over time. The two modules are concatenated by the lateral connections. Additionally, an attention strategy is adopted to adjust the weight of each factor to improve the prediction performance. An entire structure of the FAT-AttNet is illustrated in Fig 2.

#### • Fixed temporal module

In the field of time-series crop yield data analysis, 1D CNNs are effective in extracting temporal features. To handle a crop yield dataset with soil data inputs, a parallel 1D CNN architecture is proposed. This architecture consists of parallel 1D CNNs, each with 4 convolutional layers and 3

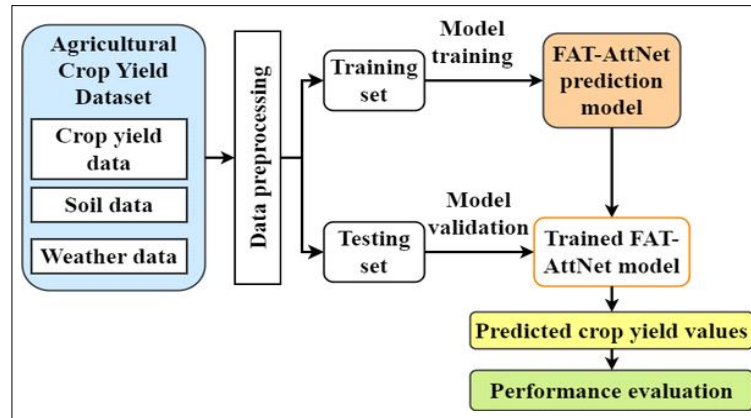


Fig 1: General layout of this research.

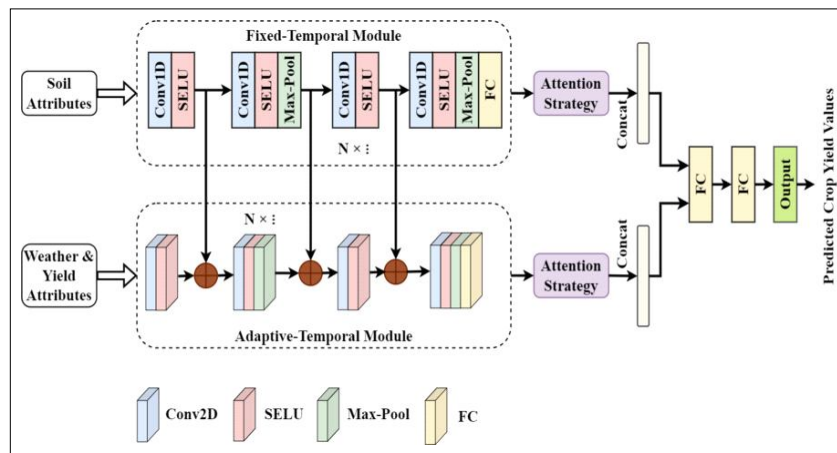


Fig 2: Structure of FAT-AttNet model.

max-pooling layers. These parallel CNNs process soil data inputs to capture their respective temporal features. The 1D CNNs use convolution kernels to convolve the time series data and extract temporal features, as defined in Eq. (2).

$$V_{ij}^* = \text{Selu} \left[ b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \omega_{ijm}^p V_{(i-1)m}^{(x+p)} \right] \quad \dots(2)$$

In Eq. (2),  $P_i$  is the dimension of 1D convolution kernel along the time dimension,  $p$  is the total time steps,  $\omega_{ijm}^p$  is the  $p^{\text{th}}$  value of the convolution kernel linked to the  $m^{\text{th}}$  feature map in the previous layer and the value of unit at  $x^{\text{th}}$  time step on the  $j^{\text{th}}$  feature map in the  $i^{\text{th}}$  layer (referred to as  $V_{ij}^*$ ),  $b_{ij}$  is the bias for this feature map,  $\text{selu}(\cdot)$  is the Scaled Exponential Linear Unit (SELU) activation function and  $V_{(i-1)m}^{(x+p)}$  is the value of unit at  $(x+p)^{\text{th}}$  time step in the  $m^{\text{th}}$  feature map of the  $(i-1)^{\text{th}}$  layer, which serves as the input to the current convolution in layer  $i$ .

#### • Adaptive temporal module

In this module, dynamic temporal relationships between  $N$  weather and  $N$  crop yield factors over time are extracted to mine further the underlying information influencing crop yield. The dynamic factors such as weather and yield data are given as input to this module. In this study,  $N$  parallel 2D CNNs are utilized to extract dynamic temporal relationships, which consist of 2D convolution, SELU, max-pooling and FC layers, as illustrated in Fig 2.

#### • Lateral connections

These connections are utilized to concatenate the outputs of two modules in each stage. These connections are added after SELU1, Pool1 and SELU2. The two modules have different information dimensions but the same number of channels, so the lateral connections concatenate the data of the two modules. The feature dimensions are  $(H, C)$  for the fixed temporal module and  $\{\alpha W, 2, C\}$  for the adaptive temporal module. The  $\{\alpha W, 2, C\}$  dimensions are reshaped and transposed to  $\{2\alpha W, C\}$ . Then, it is combined to the adaptive module.

#### • Attention strategy

In this study, the attention strategy is adopted to allocate larger weights to the crop yield, weather and soil data with a high contribution to prediction. The global spatial information is squeezed into a feature descriptor symbol and the feature-wise statistics are produced by global mean pooling. Each data creates a statistic  $z \in \mathbb{R}^c$  by shrinking  $U$  such as the  $c^{\text{th}}$  attribute of  $z$  is computed by:

$$z_c = F_{sq}(u_c) = \frac{1}{N} \sum_{i=1}^N u_c(i) \quad \dots(3)$$

In Eq. (3),  $F_{sq}(\cdot)$  is the Squeeze operation and  $u_c$  is the attribute. A nonlinear correlation between the attributes is learned to generate weights for all data attributes. Considering the complexity and generalizability of the model, 2 FC layers and 2 sigmoid functions are set to quantify the significance of each data, defined as:

$$S = F_{ex}(Z, W) = \sigma[g(Z, W)] = \sigma(W_2 \delta(W_1 Z)) \quad \dots(4)$$

In Eq. (4),  $F_{ex}(\cdot)$  is the excitation operation,  $\sigma(x)$  and  $\delta(x)$  denote the Sigmoid and Rectified Linear Unit (ReLU) activation function,  $W_1$ ,  $W_2$  represent the parameters of dimensionality-reducing and dimensionality-increasing layers, respectively. The input feature is multiplied by the corresponding weight to get a final result as:

$$\tilde{x}_c = F_{scale}(u_c, S_c) = S_c \cdot u_c \quad \dots(5)$$

In Eq. (5),  $\tilde{X} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_c]$  and  $F_{scale}(u_c, S_c)$  is the feature-wise multiplication between the scalar and the feature map  $u_c \in \mathbb{R}^c$ .

#### • FC layer

It is utilized to learn the spatial characteristics of multiple channels. In this study, the results from  $N$  parallel networks are flattened and combined into a 1D vector, connecting the output layers through two FC layers. In predicting crop yield, the output layer comprises  $N$  neurons to make predictions.

The FAT-AttNet model is designed for CYP, with its efficiency assessed through a comparison of expected values against real crop yield data.

**Algorithm 1:** FAT-AttNet model for CYP.

**Input:** Agricultural crop yield dataset with soil, weather and yield data for different crops.

**Output:** Predicted crop yield values.

1. Start.
2. Standardize every input data by Eq. (1).
3. //Train the FAT-AttNet model.
4. Use the fixed temporal and adaptive temporal modules to mine fixed and dynamic temporal correlations between different factors.
5. Apply attention strategy to assign weights for each attribute.
6. Concatenate outputs from fixed and adaptive temporal modules.
7. Apply FC layers to make final crop yield predictions.
8. Utilize test dataset to validate the model efficiency in crop yield prediction.
9. End.

## RESULTS AND DISCUSSION

This part assesses the performance of the FAT-AttNet framework with comparison to conventional techniques.

A set of existing models are utilized for the assessment of the FAT-AttNet model, including IOF-LSTM (Bhimavarapu *et al.*, 2023), CNN-BiLSTM (Saini *et al.*, 2023), MFA-BiLSTM (Krishna *et al.*, 2024) and STACNN-BiLSTM (Saravanan and Bhagavathiappan, 2024). The parameter settings for training different models are given in Table 1.

#### Evaluation metrics

**MAE:** It represents the mean absolute difference among expected and actual values.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad \dots(6)$$

In Eq. (6),  $n$  is the number of observations, such as 145 samples;  $y_i$  and  $\hat{y}_i$  are the actual and predicted values.

**MSE:** It quantifies the average squared difference between predicted and actual values.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \dots(7)$$

**RMSE:** It is the square root of MSE, indicating the average magnitude of prediction error.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad \dots(8)$$

**Correlation coefficient (r):** It measures the strength of association between actual and predicted crop yields.

$$r = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad \dots(9)$$

In Eq. (9),  $\bar{y}_i$  is the mean of the actual crop yield values.

**Mean absolute percentage error (MAPE):** It determines the mean percentage error between predicted and actual values.

$$MAPE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \times 100 \quad \dots(10)$$

**Normalized RMSE (NRMSE):** It is the normalized value of RMSE, which signifies the relative magnitude of prediction error with respect to the data mean value.

The time series comparisons of the FAT-AttNet model against existing models for CYP are displayed in Fig 3. From these analyses, it is observed that the FAT-AttNet model is the closest to the actual crop yield data. As a result, the FAT-AttNet achieved a higher efficiency for predicting different crop yields.

Fig 4 compares evaluation metrics of various CYP models for groundnut yield prediction in Tamil Nadu. The proposed FAT-AttNet demonstrates superior performance by achieving lower MAE, MSE and RMSE and a higher  $r$  compared to existing models. The FAT-AttNet attained an MAE of 0.0513, an MSE of 0.0469, an RMSE of 0.2166, an NRMSE of 0.24 and an  $r$  of 0.8617, outperforming the IOF-LSTM, CNN-BiLSTM, MFA-BiLSTM and STACNN-BiLSTM models.

Fig 5 presents the evaluation metrics for maize yield prediction using various CYP models. The proposed FAT-AttNet achieves a lower MAE (0.0586) compared to other models. It also consistently reduces MSE (0.0719), RMSE (0.2681) and NRMSE (0.295) values, along with a higher  $r$  (0.8459), demonstrating its effectiveness across all evaluation criteria. Fig 6 compares the evaluation metrics of various CYP models for moong yield prediction. The proposed FAT-AttNet demonstrates superior performance, achieving a significantly lower MAE (0.0612) compared to other models. It also consistently records reduced MSE (0.06), RMSE (0.2449) and NRMSE (0.269) values and a higher  $r$  (0.8644), highlighting its effectiveness in yield prediction.

Fig 7 illustrates the evaluation metrics of various CYP models when predicting rice yields. It is noted that the FAT-

AttNet achieved higher performance compared to the others in CYP. The FAT-AttNet achieved an MAE of 0.0576, an MSE of 0.0552, an RMSE of 0.2349, an NRMSE of 0.259 and an  $r$  of 0.86 for rice yield prediction. Fig 8 demonstrates the evaluation metrics for different CYP models in predicting urad yields. It is noticed that the FAT-AttNet attained higher efficiency in predicting crop yields due to the extraction of temporal correlations between fixed and adaptive factors related to crop yield. The FAT-AttNet achieved an MAE of 0.0741, an MSE of 0.0709, an RMSE of 0.2663, an NRMSE of 0.293 and an  $r$  of 0.8587 for urad yield prediction.

Fig 9 shows the MAPE results for predicting different crop yields using proposed and existing models. It is observed that the proposed FAT-AttNet model reduced MAPE significantly ranging between 6.1% and 9.3% for five different crops compared to the other models. Based on the above evaluations, the FAT-AttNet model demonstrates

**Table 1:** Parameter settings for training models.

| Parameters                                                               | Range   |
|--------------------------------------------------------------------------|---------|
| <b>IOF-LSTM, CNN-BiLSTM, MFA-BiLSTM and STACNN-BiLSTM and FAT-AttNet</b> |         |
| Batch size                                                               | 64      |
| Learning rate                                                            | 0.001   |
| No. of epochs                                                            | 100     |
| Optimizer                                                                | Adam    |
| Dropout                                                                  | 0.5     |
| <b>IOF-LSTM</b>                                                          |         |
| No. of hidden layers                                                     | 11      |
| No. of neurons in each hidden layer                                      | 50      |
| No. of dense layers                                                      | 4       |
| Weight decay                                                             | 0.0001  |
| <b>CNN-BiLSTM and STACNN-BiLSTM</b>                                      |         |
| No. of convolutional layer                                               | 3       |
| BiLSTM unit size                                                         | 200     |
| No. of hidden layers                                                     | 7       |
| No. of filters                                                           | 8       |
| Filter size                                                              | 2       |
| Activation function                                                      | Sigmoid |
| <b>MFA-BiLSTM</b>                                                        |         |
| Female attraction coefficient                                            | 1.5     |
| Gravity coefficient                                                      | 0.8     |
| Male cognitive coefficient                                               | 1.5     |
| Male social coefficient                                                  | 1.5     |
| Maximum iteration                                                        | 500     |
| Mutation rate                                                            | 0.05    |
| No. of offspring                                                         | 20      |
| Nuptial coefficient                                                      | 0.02    |
| Population size                                                          | 100     |
| Visibility coefficient                                                   | 0.7     |
| <b>Proposed FAT-AttNet</b>                                               |         |
| No. of 1D CNNs                                                           | 18      |
| No. of 2D CNNs                                                           | 16      |
| Kernel size                                                              | 3×3     |
| Activation function                                                      | SELU    |
| No. of FC layers                                                         | 2       |

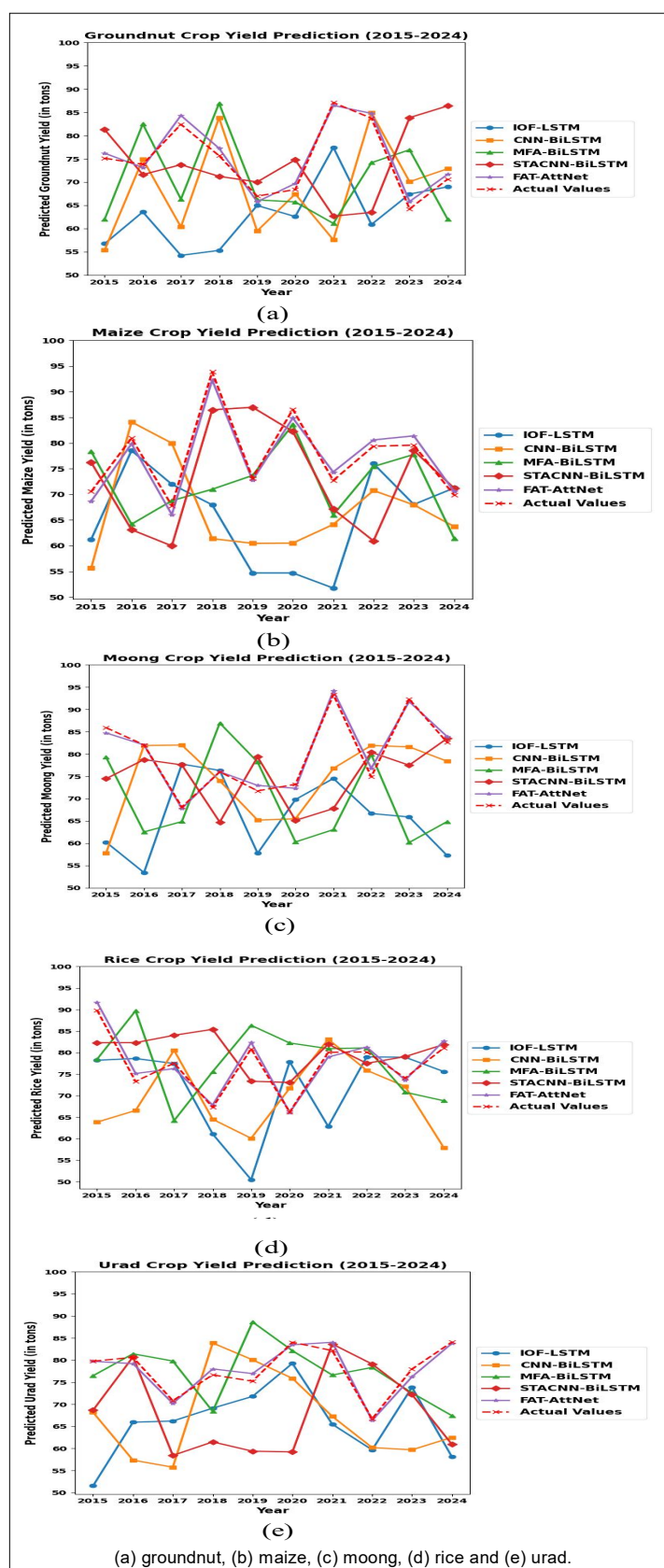


Fig 3: Comparison of proposed and existing models for CYP (in tons) from 2015 to 2024.

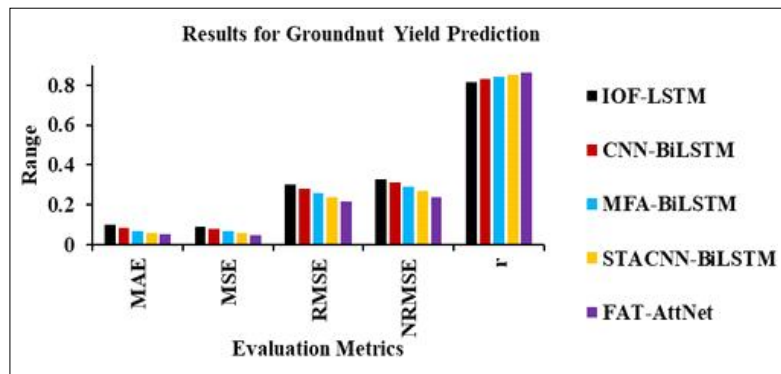


Fig 4: Comparison of evaluation metrics for groundnut yield prediction using different models from 2015 to 2024.

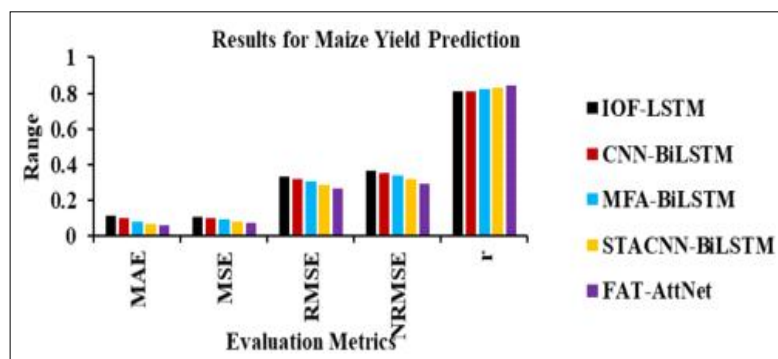


Fig 5: Comparison of evaluation metrics for maize yield prediction using different models from 2015 to 2024.

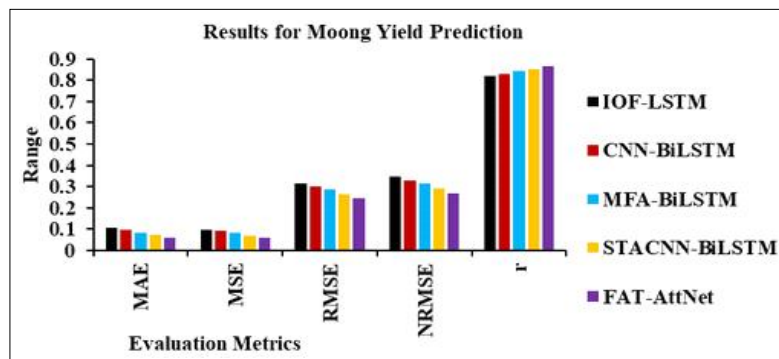


Fig 6: Comparison of evaluation metrics for moong yield prediction using different models from 2002 to 2024.

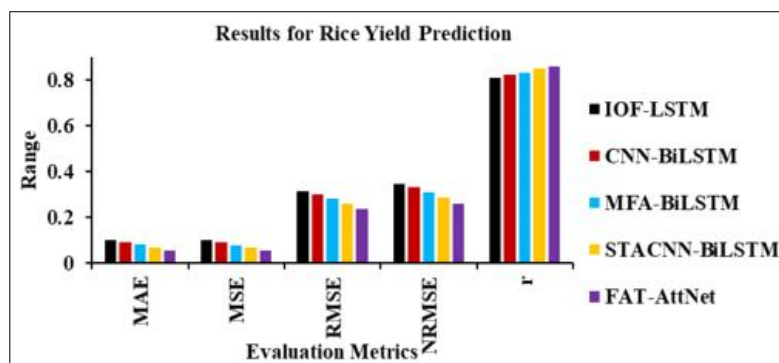


Fig 7: Comparison of evaluation metrics for rice yield prediction using different models from 2002 to 2024.

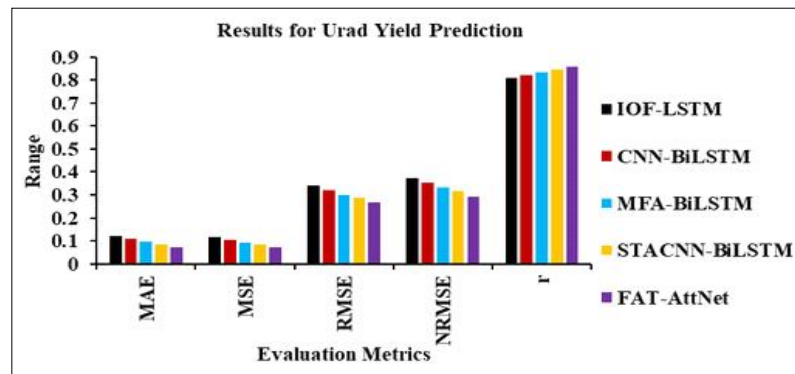


Fig 8: Comparison of evaluation metrics for urad yield prediction using different models from 2002 to 2024.

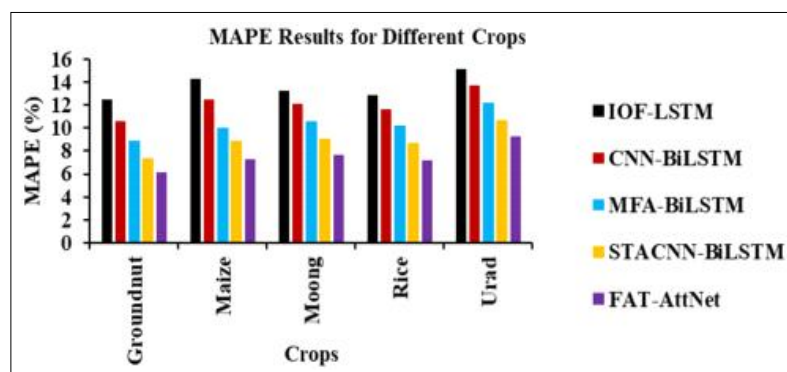


Fig 9: Comparison of MAPE for different crops.

superior accuracy in forecasting various crop yields. Its effectiveness lies in its capability to learn both static and dynamic temporal patterns between environmental factors and yield over time. As a result, this model holds practical value for farmers by enabling early yield estimation using climatic and soil-related information.

## CONCLUSION

This research presented the FAT-AttNet technique for CYP in Tamil Nadu. This model was trained using the agricultural crop yield dataset, which includes historical soil, weather and yield data. The FAT-AttNet model was constructed based on the fixed and adaptive temporal modules to capture temporal correlations between soil, weather and yield data attributes over fixed and varying times. The results of these modules were fused and weighed by the attention strategy to highlight the significant features for CYP. Following this, the FC layer was used to predict the final yield values for different crops. Finally, the performance of this FAT-AttNet model was tested against existing models for predicting groundnut, maize, moong, rice and urad crops in Tamil Nadu. The results show a superior performance of the proposed model compared to the existing models for groundnut, maize, moong, rice and urad crops. In all types of crops, IOF-LSTM, CNN-BiLSTM, MFA-BiLSTM and STACNN models had minimum MAE values of between 0.0608 and 0.0855, with the highest  $r$  of between 0.8311 and 0.85. These performance indicators are significantly lower when compared to the performance of

the suggested FAT-AttNet, which achieved the lowest MAE values ranging between 0.0513 and 0.0741, as well as the highest correlation coefficients ( $r$ ) ranging between 0.8459 and 0.8644 among all five crops. It also performed well in other metrics, indicating a higher level of robustness and reliability for the model. However, the model's performance depends on setting the appropriate hyperparameter values. Fixed parameters can impact model training for real-time CYP. Therefore, future work will focus on introducing a bio-inspired algorithm to choose optimal hyperparameters, which significantly improves the model performance.

## 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

- Al-Adhaileh, M.H. and Aldhyani, T.H. (2022). Artificial intelligence framework for modeling and predicting crop yield to enhance food security in Saudi Arabia. *Peer J. Computer Science*. **8**: e1104. doi: 10.7717/peerj-cs.1104.
- Batool, D., Shahbaz, M., Asif, H.S., Shaukat, K., Alam, T.M., Hameed, I.A., Luo, S. (2022). A hybrid approach to tea crop yield prediction using simulation models and machine learning. *Plants*. **11**: 1925. doi: 10.3390/plants11151925.

- Bhimavarapu, U., Battineni, G., Chintalapudi, N. (2023). Improved optimization algorithm in LSTM to predict crop yield. *Computers*. **12**: 10. doi : 10.3390/computers12010010.
- Dwivedi, A.K., Upreti, H., Ojha, C.S.P. (2022). Wheat yield modelling using infocrop and DSSAT crop simulation models. *Indian Journal of Agricultural Research*. **56(6)**: 646-652. doi: 10.18805/IJARE.A-5981.
- Jena, P.R., Majhi, B., Kalli, R., Majhi, R. (2023). Prediction of crop yield using climate variables in the south-western province of India: A functional artificial neural network modeling (FLANN) approach. *Environment, Development and Sustainability*. **25**: 11033-11056.
- Kalmani, V.H., Dharwadkar, N.V., 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.
- Khan, N., Kamaruddin, M.A., Sheikh, U.U., Zawawi, M.H., Yusup, Y., Bakht, M.P., Mohamed Noor, N. (2022). Prediction of oil palm yield using machine learning in the perspective of fluctuating weather and soil moisture conditions: Evaluation of a generic workflow. *Plants*. **11**: 1-19.
- Krishna, M.V., Swaroopa, K., SwarnaLatha, G., Yasaswani, V. (2024). Crop yield prediction in India based on Mayfly optimization empowered attention-bi-directional long short-term memory (LSTM). *Multimedia Tools and Applications*. **83**: 29841-29858.
- Lin, Y., Li, S., Duan, S., Ye, Y., Li, B., Li, G., Liu, J. (2023). Methodological evolution of potato yield prediction: A comprehensive review. *Frontiers in Plant Science*. **14**: 1214006. doi: 10.3389/fpls.2023.1214006.
- Luo, L., Sun, S., Xue, J., Gao, Z., Zhao, J., Yin, Y., Luan, X. (2023). Crop yield estimation based on assimilation of crop models and remote sensing data: A systematic evaluation. *Agricultural Systems*. **210**: 103711. doi: 10.1016/j.agsy.2023.103711.
- Oikonomidis, A., Catal, C., Kassahun, A. (2022). Hybrid deep learning-based models for crop yield prediction. *Applied Artificial Intelligence*. **36**: 2031822. doi: 10.1080/08839514.2022.2031823.
- Paudel, D., de Wit, A., Boogaard, H., Marcos, D., Osinga, S., Athanasiadis, I.N. (2023). Interpretability of deep learning models for crop yield forecasting. *Computers and Electronics in Agriculture*. **206**: 107663. doi: 10.1016/j.compag.2023.107663.
- Rashid, M., Bari, B.S., Yusup, Y., Kamaruddin, M.A., Khan, N. (2021). A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction. *IEEE Access*. **9**: 63406-63439.
- Saini, P., Nagpal, B., Garg, P., Kumar, S. (2023). CNN-BI-LSTM-CYP: A deep learning approach for sugarcane yield prediction. *Sustainable Energy Technologies and Assessments*. **57**: 103263. doi: 10.1016/j.seta.2023.103263.
- Saravanan, K.S. and Bhagavathiappan, V. (2024). Prediction of crop yield in India using machine learning and hybrid deep learning models. *Acta Geophysica*. pp 1-20. doi: 10.1007/s11600-024-01312-8.
- Seireg, H.R., Omar, Y.M., Abd El-Samie, F.E., El-Fishawy, A.S., Elmahalawy, A. (2022). Ensemble machine learning techniques using computer simulation data for wild blueberry yield prediction. *IEEE Access*. **10**: 64671-64687.
- Shaikh, T.A., Rasool, T., Lone, F.R. (2022). Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Computers and Electronics in Agriculture*. **198**: 107119. doi: 10.1016/j.compag.2022.107119.
- Srivastava, A.K., Safaei, N., Khaki, S., Lopez, G., Zeng, W., Ewert, F., Rahimi, J. (2022). Winter wheat yield prediction using convolutional neural networks from environmental and phenological data. *Scientific Reports*. **12**: 3215. doi: 10.1038/s41598-022-06249-w.
- Talaei, K.T., Ould Slimane, H., Kaabouch, N. (2023). Deep learning: Systematic review, models, challenges and research directions. *Neural Computing and Applications*. **35**: 23103-23124.
- Thimmegowda, M.N., Manjunatha, M.H., Sathisha, G.S., Huggi, L. (2024). Weather based yield prediction of Pigeonpea crop with the resource use efficiency under varied sowing windows. *Legume Research-An International Journal*. 1-8. doi: 10.18805/LR-5308.
- Trentin, C., Ampatzidis, Y., Lacerda, C., Shiratsuchi, L. (2024). Tree crop yield estimation and prediction using remote sensing and machine learning: A systematic review. *Smart Agricultural Technology*. **9**: 100556. doi: 10.1016/j.atech.2024.100556.
- Xia, L., Robock, A., Scherrer, K., Harrison, C.S., Bodirsky, B.L., Weindl, I., Heneghan, R. (2022). Global food insecurity and famine from reduced crop, marine fishery and livestock production due to climate disruption from nuclear war soot injection. *Nature Food*. **3**: 586-596.
- Zvobgo, L., Johnston, P., Olagbegi, O.M., Simpson, N.P., Trisos, C.H. (2023). Role of indigenous and local knowledge in seasonal forecasts and climate adaptation: A case study of smallholder farmers in Chiredzi, Zimbabwe. *Environmental Science and Policy*. **145**: 13-28.