Sustainable Energy Technologies and Assessments 57 (2023) 103263

Contents lists available at ScienceDirect

Sustainable Energy Technologies and Assessments

journal homepage: www.elsevier.com/locate/seta

CNN-BI-LSTM-CYP: A deep learning approach for sugarcane yield prediction

Preeti Saini [a], Bharti Nagpal [b], Puneet Garg [c], Sachin Kumar [d,*]

[a] NSUT East Campus (Formerly AIACTR), USICT, Guru Gobind Singh Indraprastha University (GGSIPU), Delhi, India
[b] Department of Computer Science and Engineering, NSUT East Campus (Formerly AIACTR), Delhi, India
[c] Department of Computer Science and Engineering, ABES Engineering College, Ghaziabad, U.P., India
[d] Big Data & Machine Learning Lab, South Ural State University, Chelyabinsk, 454080, Russian Federation

A R T I C L E I N F O

Keywords:
Forecasting
Yield
Prediction
ML-GPR
CNN-Bi-LSTM
ARIMA
GRU

A B S T R A C T

Sugarcane (Saccharum officinarum L.) is one of the principal origins of sugar and is also known as the main cash crop of India. About 19.07% of the total production of the world's sugar requirement is fulfilled by India. Traditionally, Statistical approaches have been utilized for Crop yield prediction, which is tedious and time-consuming. In this direction, the present work proposed a novel hybrid CNN-Bi-LSTM_CYP deep learning-based approach that includes convolutional layers to extract the relevant spatial information in a sequence to Bi-LSTM layers that recognize the Phenological long-term and short-term bidirectional dependencies in the dataset to predict the Sugarcane crop yield. The experimentation was performed and validated on the historical dataset from 1950 to 2019 years of the major Sugarcane-producing states of India. The preliminary results shown that the CNN-Bi-LSTM_CYP method performed well (RMSE:4.05, MSE:16.40) in comparison to traditional Stacked-LSTM (RMSE:8.8, MSE:77.79), ARIMA (RMSE:5.9, MSE:34.80), GPR (RMSE:10.1, MSE:103.3), and Holt-winter Time-series (RMSE:9.9, MSE:99.7) techniques. The study concluded that the predicted sugar yield has a minimal relative error concerning the ground truth data for the CNN-Bi-LSTM_CYP approach proving the proposed model's efficiency.

Introduction

The Food and Agriculture Organization statistics signified that around 79% of the sugar is produced by processing sugarcane crops worldwide. Approximately 40% of the sugar requirement is fulfilled by Brazil, followed by India. From 2021 to 2022, about 181.18 million metric tonnes of sugar were produced, and sugar consumption exceeded 175 million metric tonnes [1]. The Population uplift of approx. 9.7 billion by 2050 will become a significant factor behind these requirement's urge worldwide. Humans will face the Consequences of this, 'Will enough food be provided at an affordable price to complete the dietary requirements?' [2,3]. Therefore, based on these perspectives, predicting the crop before pre-harvest has become an important research domain worldwide [4]. There is a direct link between the production of Crops and their price in the market, and if there is a decrement in production can lead to a massive price hike. To overcome this demand–supply gap, there is a continuous requirement to forecast the situation of Cash crops to enhance the country's export business to a reasonable extent [5]. Several Crop Yield estimation models were proposed in the literature based on crop-related that provided reasonable outcomes [6–8]. With the emergence of the artificial intelligence field, deep learning methods, a subset of Machine learning, have overcome the limitation of traditional approaches. Machine learning (ML) and Deep learning (DL) techniques have been extensively utilized for crop yield prediction due to their ability to recognize non-linear patterns in a vast dataset [9,10]. In this study, we built a novel CNN and Bi-LSTM hybrid deep-learning architecture for crop yield estimation of the sugarcane crop. We performed the experimentation for the study regions of India. Sugarcane Crop is one of the major cash crops of India. In 2021, approximately seventeen lakh tonnes of sugar were exported outside the country; in 2020, this amount was 4.5 lakh tonnes [11]. The latest information for the 2021 year presented that there were 479 sugar mills exists in the country that produced 77.9 tonnes of sugar [12,13]. Therefore, in this direction, the assessment of yield and production of this crop plays a crucial role in decision-making policies for export by the farmers [6,14].

* Corresponding author.
E-mail addresses: preeti.saini@dseu.ac.in (P. Saini), bharti.nagpal@nsut.ac.in (B. Nagpal), puneetgarg.er@gmail.com (P. Garg), kumars@susu.ru (S. Kumar).

https://doi.org/10.1016/j.seta.2023.103263
Received 13 September 2022; Received in revised form 24 April 2023; Accepted 28 April 2023

ARIMA linear model was popular previously for time series forecasting of crops [6,14–17]. Nowadays, Deep Learning techniques overcome the Linear model issues as they cannot capture the non-linear patterns in the dataset [18–20]. Several studies signify the performance of the ARIMA model that had provided good results for linear time series forecasting and prediction. In a study, Vishwajith et al. [6] forecasted sugarcane production in India by 2020 year by considering the Area, Yield, and productivity as variables of the crop from 1951 to the 2012 year and found that the Auto-Regressive Moving average (2,1,1) model was appropriate for sugar crop experimentation. A similar approach was presented by Mishra et al. for a study of the production of Sugar crops in India using the ARIMA method. The results were evaluated using RMSE and MAPE metrics for the dataset from 1950 to 2015 [7]. The emergence of Machine learning and Deep learning techniques with remote sensing revolutionized the field of agriculture [20–24]. They have the capability of identification of complex relationships that exist in the patterns of the dataset among various parameters that can influence the growth of crops [25–28]. In the study, Haider et al. utilized a variant of Recurrent Neural Network, i.e., LSTM (Long Short-Term Memory), an approach of Deep Learning to forecast wheat production by improvising the pre-processing phase using a smoothing function. They compared the results of RNN with the ARIMA model and found that LSTM outperformed with an R^2 value of 0.81 [29]. A similar technique of DL was used for soil moisture forecasting based on humidity and climatic factors to forecast the soil moisture in Thailand, a DNN-LSTM method was applied over the irrigation dataset that was collected with smart sensors, and the outcomes of the proposed study produced a 0.72 percent RMSE value. However, this approach cannot consecutively check the learning rate of the Adam optimizer that increased the timings of the training and testing phase of the model [30,31]. A similar approach of LSTM was utilized in applying the recurrent neural network method to predict wheat crop yield in the Punjab region of India. The results were evaluated as RMSE 147.12 and MAE as 60.50 [32]. Many authors addressed the capabilities of machine learning methods such as K-means, regression model, and Naïve Bayes analysis on weather datasets to employ Regression Techniques to carry out prediction tasks with an RMSE value of 1.06 across major districts of India [33] and the comparison among ANN and regression model was discussed [34]. In this direction, Ansarifar et al. [8] applied a linear regression model for forecasting soya bean and corn crop yield by considering genotype data. They exhibited that the proposed regression model achieved 8% lower RMSE in three experimental sites and suggested hyperparameter tuning can be resolved in future experimentation. Prior to the Deep learning technique, Linear models, and Machine learning method were highly used despite their various issues [35,36,50]. Similarly, the study was carried out using ML to predict crop yield in the Tamil Nadu region of India using ANN techniques and found that the Random Forest method performed well in predicting accurate results by using the area and temperature information of the region [37].

The literature shown that most of the existing research in context of crop prediction were based on statistical and linear models. The existing linear models are computationally expensive as, in every iteration the method needs to determine the differencing values that enhance the training time of the method. These models do not capture the non-linear relationship among the crop attributes, and therefore, not suitable for long-term forecasting of time series data. Another constraint is to improve the training and testing time taken by machine learning and linear models to handle the large crop datasets. Recently, LSTM technique have also gained popularity in prediction and forecasting of long-term dependencies but they suffer from the Vanishing gradient issue. In this context, our work aims to propose a novel hybrid deep learning-based CNN with Bi-LSTM network that can cope up with these gaps.

Materials and methods

Study area

In this article, we focussed on one of India's major cash crops, i.e., the Sugarcane Crop. The investigation includes an analysis of the sugarcane crop in India's five major contributor states (Uttar Pradesh, Maharashtra, Karnataka, Tamil Nadu, and Bihar). As per statistical information for the 2021 year, Uttar Pradesh is India's main contributor state of the Sugarcane Crop (approx. 48%). The description of the sugarcane crop growing and harvest period in the study regions is shown in Table S1.

Proposed Hybrid Approach (CNN-Bi-LSTM_CYP)

This section presents the methodology that discusses the hybrid deep learning architecture and working of the CNN-Bi-LSTM-CYP Model. A solution is built to overcome the limitations that exist in the literature to predict crop yield. The proposed approach used in this study is shown in Fig. 1.

The proposed hybrid architecture of deep learning utilizes multiple layers in the network to investigate the main features in the dataset. Many authors suggested that the model's performance can be enhanced effectively by increasing the depth of neural networks as deeper features are extracted and learned by the architectures [38,39]. The investigation includes utilizing the CNN and Bi-LSTM deep neural networks for yield prediction. The Primary purpose of Convolutional Layers is to extract relevant features and learn from the Sugarcane_CYP time series dataset to determine the yield pattern. Then Bi-LSTM network finds the short-term and long-term dependencies in the dataset in both directions. A Convolutional Neural Network (CNN), known as 'ConvNet,' facilitates feature learning through various processing layers. The structure of CNN consists of one input layer, several hidden layers, a fully connected layer, and one output layer [40–43].

The Bi-LSTM network combines two LSTM units that take input from both the forward and backward directions. To illustrate the working of Bi-LSTM, a recurrent neural network (RNN) needs to understand. An RNN is recognized for "memory" because its intakes information from earlier contributions is considered suitable for time series or sequence data [44]. The structure of RNN mainly comprises the input layer, hidden layer & output layer. Every part maps each input through the hidden state to evaluate yield sequentially, utilizing the concept of the Back Propagation algorithm by computing errors from different layers to the yield layer. To overcome the problems arising from RNN, a long short-term memory (LSTM) was introduced, which supports long-duration memory capacity to store elements and support long-term dependencies [45,46].

The structure of the Bi-LSTM technique is arranged in two RNNs aligned independently, as shown in Fig. 2. They enhance the accuracy of the prediction model as they process the information from front to back and vice versa [47–49].

In the proposed network, initially, the dataset is passed through the pooling and convolutional layers that act as a pre-processing layer to extract the relevant information from the data. This meaningful data is then used in the hybrid neural network. The convolutional layers applied the convolution operator between the convolutional kernel and input to create the new set of features. This procedure results in the matrix format, which consists of feature values. Several Convolutional kernels are applied over the input to get the convolved features that contain more meaningful data than the neural network input. The outcomes of max pooling layers are then fed to the Bi-LSTM network that applies the three gates to filter the further information from the data discussed in the subsequent section based on Eqs. (1) to (8). The output of these gates becomes an input to the fully connected layer that connects each piece of information to provide the results. Finally, a relu activation function is applied to assign the features to the classes to

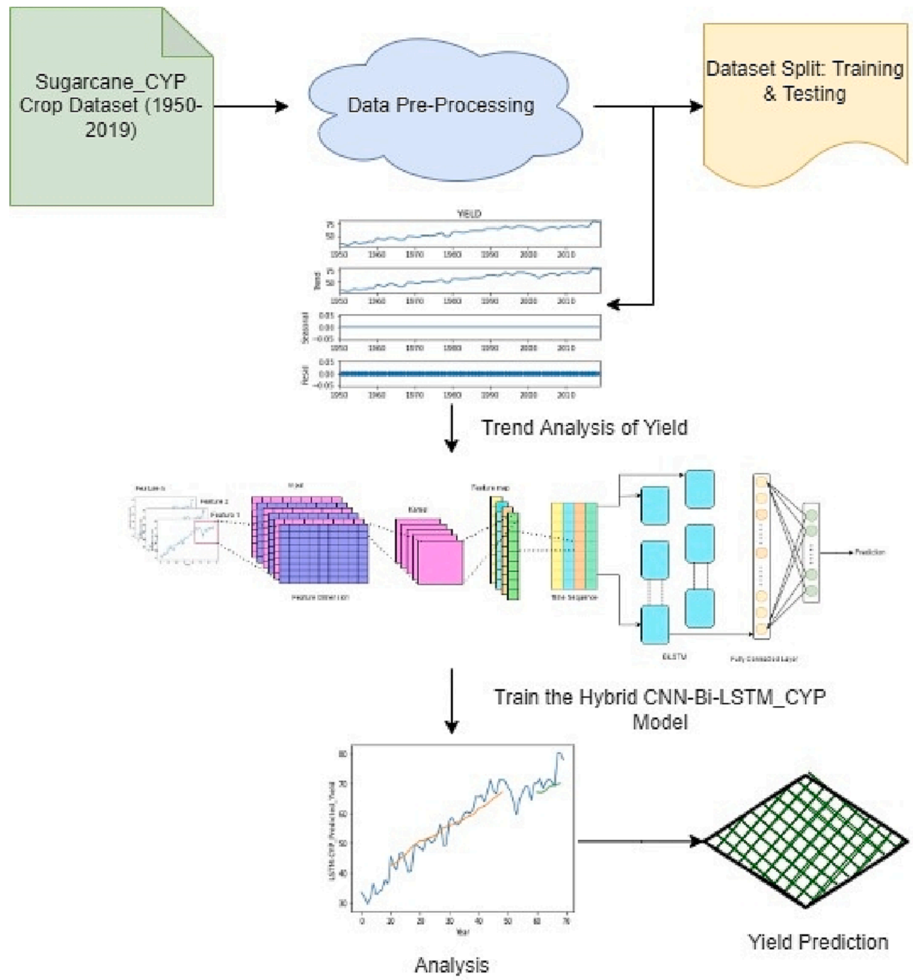


Fig. 1. Methodology of Proposed Hybrid CNN-Bi-LSTM_CYP Model.

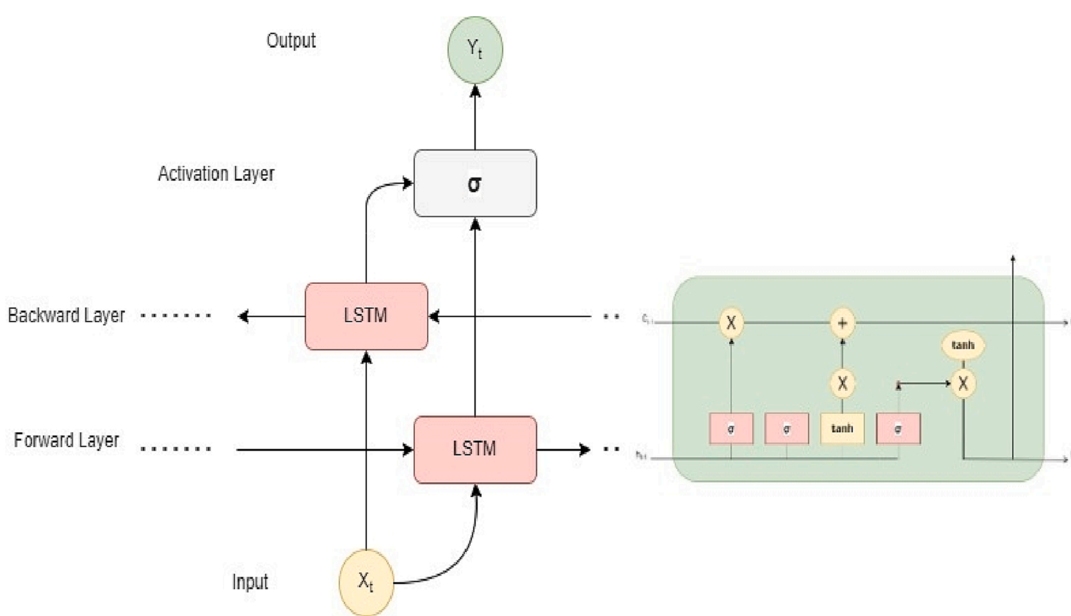


Fig. 2. Structure of Bi-LSTM Technique.

produce the desired outcome. The processing of the prediction model mainly includes three steps: Data pre-processing, Convolution operation, and Bi-LSTM, a fully connected layer, which can be explained with Eqs. (1) to (8).

The forget gate deals with what to omit from the cell and can be represented by Equation (1).

$$f_t = \sigma\left(w_f\left[h_{t-1}x_t\right] + b_f\right) \tag{1}$$

Where x_t is the current input, h_{t-1} is the hidden state of the last step, b is the bias, and w is the weight of the matrices.

In the following, the input gate utilized tanh and sigmoid functions to refine the output stages of the cell, represented via Eq. (2) to Eq. (6).

$$\widehat{c}_t = \tanh(w_c[h_{t-1}x_t]+b_c \tag{2}$$

Where $\widehat{c}_t$ is the cell state, h_{t-1} is the hidden state of the last step, b is the bias, and w is the weight of the matrices.

The sigmoid layer creates a filter as per Eq. (3),

$$\mathrm{U}_t = \sigma(w_U[h_{t-1}x_t] + bu \tag{3}$$

After processing throughout the input gate, the sigmoid layer filters the state provided as output (O), and all cell states passed through the function to be in the range of −1 to 1.

$$o_t = \sigma(w_0[h_{t-1}x_t] + b_0 \tag{4}$$

In the final step, a hidden state (h_{t-1}) is computed by the multiplication of scale cell state ‘C’ and filter ‘O’ value to pass to the next cell, and this can be shown in Eq. (6).

$$h_t = O_t{*}\tanh(C_t) \tag{5}$$

The point-wise multiplication is computed between the result of two layers to produce the memory’s ht.

In Bidirectional LSTM used, both the forward and backward hidden state of two LSTM at the same position can be presented in Eq. (7) & Eq. (8).

$$\overrightarrow{h_t} = \overrightarrow{LSTM}(C_{t-1}, h_{t-1}) \tag{6}$$

$$\overleftarrow{h_t} = \overleftarrow{L}\,STM(C_{t-1}, h_{t-1}) \tag{7}$$

$$h_t = \mathrm{LSTM}\left(\overrightarrow{h_t}, \overleftarrow{h_t}\right) \tag{8}$$

Here, h_t is the concatenated outcome of forwarding and backward LSTM to determine long-term and short-term dependencies.

The proposed Algorithm to train the CNN-Bi-LSTM_CYP network is illustrated as follows:

Algorithm: Train the CNN-Bi-LSTM_CYP-based DNN Network

```
   Input: Historical Sugarcane_CYP Dataset
   Output: Prediction and Forecasting of Yield
1  Sugarcane_CYP: Major Sugarcane producer state's yield collection
2  Initialization of Weights (w) and Parameters {x, y, n, i}
3  While the termination criteria are not met do
4      X← {Sugarcane_CYP_Features}. Valuesy ←{Sugarcane_CYP_Classes}. Values
5      Train_CYP_Data, Test_CYP_Data, Valid_CYP_Data ← Train_Test_Split (x, y,
       0.72,0.28)Batch_Size ← 64
6      While(n < epochs) do
7          CNN-Bi-LSTM_CYP ← Sequential_Model
            ([
            Embed_Layers (Train_CYP_Data.length,
            Result.length, Train_CYP_Data.columns),
            CYP_Convolutional_Layers (filters = 32, kernel_size = 3,
              strides = 1, padding="causal",
              activation="relu",
              input_shape = [Timesteps, feature_count], Result.length)
            CYP_Bi-LSTM_Layers (32, return_sequences = True, Result.length)
            CYP_Bi-LSTM_Layers (32, return_sequences = True, Result.length)
             Dense_Layer (Result.length, activation='relu'))
            ])
```

(continued on next column)

(continued)

Algorithm: Train the CNN-Bi-LSTM_CYP-based DNN Network

```
8              Loss ←'Mean square error,' optimizer ← Adaptive Moment Estimation'
9              CNN-Bi-LSTM_CYP Model. Compile (Loss, optimizer, metrics =
               ["mse"])
               CNN-Bi-LSTM_CYP Model. Train (Train_CYP_Data, Batch_Size, epochs,
               Valid_CYP_Data)
10         End
11     End
```

Dataset

The Present study focused on India’s major Sugarcane producer states, namely Uttar Pradesh, Maharashtra, Karnataka, Tamil Nadu, and Bihar. The experimentation was performed using historical data of the last 70 years of India, collected from the Directorate of Economics and Statistics under the Ministry of Agriculture, India [13]. State-wise statistics refer to the Statistical Abstract of a particular state. The Statistical state data was available from 1990 to 2019 years. The parameters exist in the dataset, such as Area, Area under irrigation, and production attributes of a specific region to compute the crop yield factor. The statistical information revealed that maximum yield was attained in the 2017–18 year with 80,198 kg/ha over the 4.74 million hectares, whereas the minimum was attained with 29,495 kg/ha in the 1952–53 year in 66.29% irrigated area in India.

Fig. S1 represents the sugar Contribution of these listed States throughout the country [11]. The highest sugar production in Uttar Pradesh was 48.35%, followed by 18.01% in Maharashtra State.

Results and discussion

The proposed approach was implemented using Python Programming in Jupyter Notebook and WEKA tool. The hardware requirements include an i5-11300H processor, an X-64-based PC, and 16 GB RAM.

The Descriptive Statistics of Sugarcane crop yield (tonnes/ha) and Production (Tonnes) from 1950 to 2019 are presented for India in Table 1. The minimum yield was recorded in 1952–53 with 29.3 (tonnes/ha), and the maximum growth pattern was observed in 2017–2018 with 80.198 (tonnes/ha). Therefore, these findings indicate the increment in Sugar production in India from 1950 to 2019. As per Table 1, there was an increment in Sugarcane Yield in Karnataka, Tamil Nadu, and Bihar with 8.4%, 1.8%, and 6.3%, respectively. The total Sugarcane mean yield in India was 56.5. The highest mean is observed in Tamil Nadu with 103.06, whereas the lowest average is in Bihar with 46.4, respectively.

The graphical representation of Sugarcane Crop yield in Fig. S2 represents the variations in the growth of sugarcane crops over the years due to climatic conditions, soil moisture, irrigation facilities, fertilization, etc., in the Uttar Pradesh, Maharashtra, Karnataka, Tamil Nadu, Bihar, and India.

Evaluation metrics

The performance of the proposed approach was evaluated using the Root Mean Squared Error (RMSE) and Mean Squared Error (MSE) computed by Eqs. (9) and (10), respectively. The RMSE evaluates the standard deviation of the surplus data items, and the MSE measures the variance of the residuary.

$$\mathrm{MSE} = \frac{1}{M}\sum_{j=1}^{M}\left(w_j - \widehat{w}\right)^2 \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{M}\sum_{j=1}^{M}\left(w_j - \widehat{w}\right)^2} \tag{10}$$

Table 1
Descriptive Statistics of Sugar Crop yield (in tonnes/ha) from 1950 to 2019.

State/Country	Mean	Standard Deviation	Skewness	Kurtosis	Minimum	Maximum
India	56.5	13.5	−0.3	−0.9	29.4	80.1
Uttar Pradesh	61.3	7.16	1.9	3.2	52.3	80.81
Maharashtra	79.5	7.7	−1.2	1.8	57.9	92.1
Karnataka	86.3	8.3	−0.14	0.87	65.8	102.9
Tamil Nadu	103.06	5.7	−0.99	1.1	87.1	111.5
Bihar	46.4	6.93	−0.94	2.2	24.6	59.2

Where, M = no. of dataset values; $\widehat{w}$ = Observed Yield; w_j = yield's mean value.

The results are represented in the following section and a discussion of the experimental setup.

Comparison with traditional methods

This section illustrates the result obtained using existing machine learning and linear models. In traditional statistical approaches, the ARIMA method was highly utilized for prediction and forecast purposes [51], but nowadays, deep neural network techniques overcome the issues of conventional forecasting methods. The ARIMA (2,1,0) was found suitable for predicting the Yield of India with an AIC of 253.7. The auto function of the ARIMA model automatically fetches the appropriate model with the smallest AIC value. Initially, data stationarity was checked by using ADF Test. The outcomes show that the input was not stationary ($p < 0.5$). Therefore, first, differencing was performed to make the data stationary to satisfy the condition that p should not exceed the 0.5 value. The model was further initialized with p, d, and q parameters in a range (0,4) and (0,2).

Further, this linear model was evaluated for all the considered state data. The computed results revealed that for the Tamil Nadu state, the ARIMA model (0,1,1) seems best with AIC = 53.6. For the Maharashtra state, found ARIMA (0,1,0) to be a good fit with the smallest AIC of 49.6. The Karnataka state yield forecasting model with ARIMA (0,1,1) with AIC = 62.2, similarly for the Uttar Pradesh State (UP), is best with 43.1 in the (1,1,0) model. The Bihar state dataset found ARIMA (0,1,0) suitable, with the lowest AIC as 54.7.

The Gaussian process regression based on Bayesian inference used 5 lag order to evaluate the results on the Sugarcane_CYP dataset. This analysis infers that the increasing trend pattern was observed in the upcoming years in the Sugarcane Yield in India by 2029. However, in Maharashtra state, there could be a decline of approx.—6% Yield from 79.5 tonnes/ha to 74.6 tonnes/ha for 2018.

Experimental analysis of proposed (CNN-Bi-LSTM_CYP) approach

The proposed approach overcomes the issues of Recurrent Neural networks and other existing methods by utilizing several hidden layers in their processing. During experimentation, split the dataset into 72% training data, and the rest was used for testing the model. The hyperparameters utilized in this study to develop the proposed model are illustrated in Table S2.

The CNN-Bi-LSTM_CYP model used several timesteps to determine the variability over the years and to predict the subsequent year's data. In this study, we used a time step of 10 years to prepare the model learned from the last ten years to predict the subsequent yield. The structure of the implemented model includes the conv1d layers in a sequence to a stacked Bi-directional LSTM layer with one dense layer. There were around 50 neurons within each DNN-LSTM Model Layer, and a 0.2 value was used as drop-out regularization to avoid overfitting issues in the network. The model was suitably trained with a batch size of 64 with 32 filters. An adaptive moment estimation optimizer was utilized to minimize the loss incurred in the model.

Validation results of the model

The results were validated concerning the ground truth data of the 2011–2019 years by computing the relative deviation between the actual ground truth data and the predicted yield of the Sugarcane Crop in India, represented in Table S3.

The graphical analysis of the observed and predicted yield of the Sugarcane Crop evaluated using the proposed model CNN-Bi-LSTM_CYP, Stacked_LSTM, ARIMA, Holt-winter time series method, ML_GPR are shown in Fig.S3. These results revealed that CNN-Bi-LSTM_CYP accurately predicts the Sugarcane yield over the period in India with Minimal errors.

The actual and observed yield results of all the state's models signified the accurate results for Maharashtra state followed by Bihar, Karnataka, and Tamil Nadu with an RMSE value at regional data analysis as 6.95, 11.63, 12.7, 16.4, respectively.

Performance evaluation

The analysis presents the rising and falling trend in the crop yield during the last 70 years, but in recent years, there seems to be an increasing trend compared to an earlier time. Table 2 illustrate the result analysis among CNN-Bi-LSTM_CYP, Stacked_LSTM, Regression, and time series methods. The ARIMA (2,1,0) was next in the proposed technique's performance. The main drawback of the Deep Neural method is that it requires a vast amount of training data. Therefore, the accuracy of the deep network majorly depends on the availability of the dataset. In our case, we trained the model with 70 years of data, and we can further enhance the accuracy of the results if we introduce the model with the daily availability of the data. The performance of the proposed work is evaluated using RMSE and MSE, explained in Table 2.

As per the tabular analysis of Table 2, the proposed model shows a 0.2% average relative deviation and a root mean square error (RMSE) value of 4.05 between the observed and actual forecast. In comparison, the ARIMA Model & GPR average relative deviation of 27.7%,48.2% calculated & 5.9, 10.1 RMSE was observed.

Comparative analysis with existing hybrid approach

In this study, we have compared the performance of the proposed approach on a publicly available dataset (https://github.com/saeedkhaki92/CNN-RNN-Yield-Prediction) implemented using the CNN-RNN approach [52]. The dataset was comprised of 25,345 instances with 395 attributes. Our proposed CNN-BI-LSTM_CYP model outperformed with a mean squared error of 3.06 and a mean absolute error of 3.52 compared to the CNN-RNN approach RMSE in the range of 4.15–4.91, represented in Table 3.

Table 2
Comparison with existing models on Sugarcane_CYP Dataset.

Prediction Model	RMSE	MSE
Stacked LSTM	8.82	77.79
Linear Regression	4.9	24.22
ARIMA	5.9	34.81
ML_GPR	10.1	103.3
Holt-Winters Time Series	9.98	99.72
Proposed Method (CNN-Bi-LSTM-CYP)	4.05	16.40

Table 3
Comparative results with other research studies.

Author	Dataset	Methodology	RMSE
Khaki et al. [52]	Publicly available data	CNN-RNN	4.15
Proposed Method (CNN-Bi-LSTM_CYP)	Publicly available data	CNN-Bi-LSTM_CYP	1.74

The results of the proposed approach indicate the novelty of our work in dealing with a large dataset by considering the proper data processing mechanism. In their article, they used the 3,50,000 iterations with 0.03% of the learning rate, whereas we used the 0.005 learning rate with 500 epochs and a Validation split of 0.05 to train the network.

Forecasting of sugarcane yield

The forecasting of sugarcane yield is presented in Table S4 for Uttar Pradesh, Maharashtra, Karnataka, Tamil Nadu, and Bihar states of India from the 2020–2029 year.

Table S4 exhibits the increasing trend pattern in the yield expectancy of Major producers of Sugarcane Crops in the subsequent years till the 2029 year. The graphical representation of Crop Yield prediction and forecast results of UP and Maharashtra regions are shown in Fig. 3a, 3b, for Karnataka and Tamil Nadu are presented in Fig. 4a, 4b, and Bihar region results are depicted in Fig. 5.

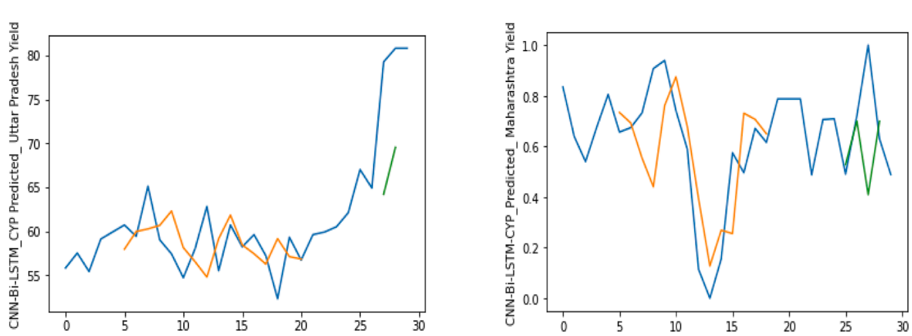
The forecasting results illustrate that in the upcoming years, India will have a rising trend in the Yield expectancy of Sugarcane Crop, which is the highest in 2029. The experimentation exhibits a decline in yield from 2018 to 19 by 2.8% and found a decline the production from 405.4 million tonnes to 355.7 million tonnes.

In Fig. 3, Fig. 4, and Fig. 5, the blue line presents the train_predict data that can be evaluated using the code snippet: (trainPredictPlot [look_back:len(train_predict) + look_back,:] = train_predict) whereas, the orange line signifies the test_predict(testPredictPlot[len(train_predict)+(look_back*2) + 1:len(df)-1,:] = test_predict) using the look-back of the last 10 years of training data. The comparison was performed among the predicted and observed yield values in a dataset, based on which forecasting can be evaluated using detailed region data in the green line. Among the discussed methods, ML-GPR has exhibited high RMSE, leading to more error in the actual to the observed value of Sugarcane Yield.

In this study, to deal with non-linear and non-stationary characteristics of crop yield state-of-the-art deep learning approaches were combined to improve the precision of prediction in handling the Sugarcane_CYP time series dataset. The current Studies for Crop yield prediction based on statistical, and linear models does not consider these characteristics of a crop simultaneously [6–8]. Therefore, we presented a CNN with Bi-LSTM based hybrid deep learning approach to take advantages of both networks in the present work. The CNN focuses to extract the relevant features and learn from the Sugarcane_CYP time series dataset to determine the yield pattern, whereas the Bi-LSTM network finds the short-term and long-term dependencies in the dataset in both directions to strengthen the information for forecasting the long-term time series data. The CNN-Bi-LSTM_CYP method improved the training & testing time taken in existing baselines models by including the relevant parameters and, overcome the vanishing gradient issue of LSTM technique, and achieved a 0.2% average relative deviation in predicted yield [32]. The experimental outcomes indicate that proposed approach provides the lowest error between the predicted and observed yield with 4.05 tonnes/ha. The proposed model leverages the advantage of feature engineering in Hybrid DL model and discards the irrelevant parameters to attain the higher accuracy in comparison to traditional machine learning and linear models. Methods like ML_GPR, LSTM, and ARIMA shown low performance in terms of accuracy and speed. The findings substantially exhibit a higher accuracy in validation in comparison to other methods. In general, CNN-Bi-LSTM_CYP method also evaluated using publicly available dataset to represent their efficiency in handling large crop datasets and achieved a better accuracy with 1.74 [52]. Furthermore, the yield estimates were upgraded from country level to major sugarcane producer regions presented in Fig. 3, Fig. 4, and Fig. 5 to evaluate the model applicability at city level, and found that ML-GPR more prone to errors in actual to observed yield while processing the dataset. The study can be applied for monitoring the vegetation growth and estimation of crop yield. The findings in this study demonstrated the applicability of deep learning framework in the field of smart farming in managing and predicting the crop yield. In addition, the potential limitation of the CNN-Bi-LSTM_CYP lies in determination of optimal hyperparameters and in future, we will work out on this problem to automate the searching of best hyperparameters to train the DL model. Another future aspect includes, the data intensive nature of the DL models can be managed by incorporating the network with more data characteristics including daily monitored data, remote sensing derived features so that the proposed model can learn better from data characteristics and further enhance the accuracy of the model.

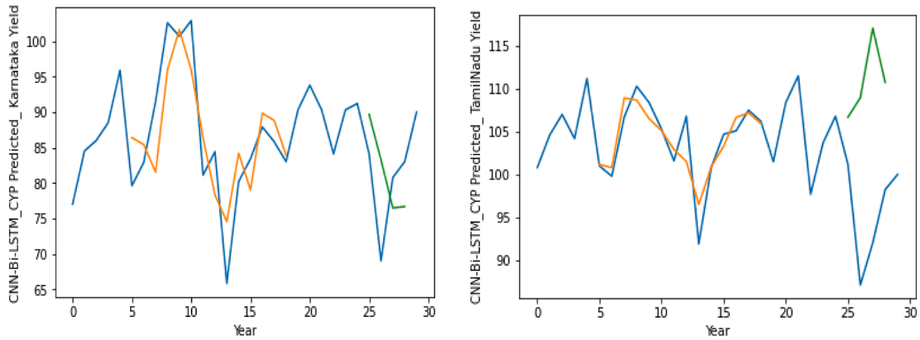
Conclusion

In this study, we investigated the various linear, machine learning, and deep learning models for crop yield prediction and built the hybrid deep learning network. The CNN-Bi-LSTM_CYP approach provides a novel mechanism to overcome the limitations of linear and regression models to analyze the statistical data for time series forecasting. The efficiency of the proposed approach was also assessed on a publicly available dataset compared to other hybrid models. We have taken advantage of both the CNN and Bi-LSTM layers to capture temporal and spatial information in the dataset. Feature engineering was performed to train the model with appropriate features. The performance of the CNN-Bi-LSTM_CYP approach was evaluated using evaluation metrics RMSE and MSE and found to be the best with the lowest RMSE of 4.05 among Stacked_LSTM, ARIMA, HoltWinter time-series, and Machine Learning-



3a) UP_Sugarcane Yield Prediction 3b) Maharashtra Sugarcane Yield Prediction

Fig. 3. Forecasting and Prediction of Sugarcane Yield of Uttar Pradesh & Maharashtra States using CNN-Bi-LSTM_CYP Model.



4a) Karnataka Sugarcane Yield Prediction 4b) Tamil Nadu Sugarcane Yield Prediction

Fig. 4. Forecasting and Prediction of Sugarcane Yield of Karnataka & Tamil Nadu States using CNN-Bi-LSTM_CYP Model.

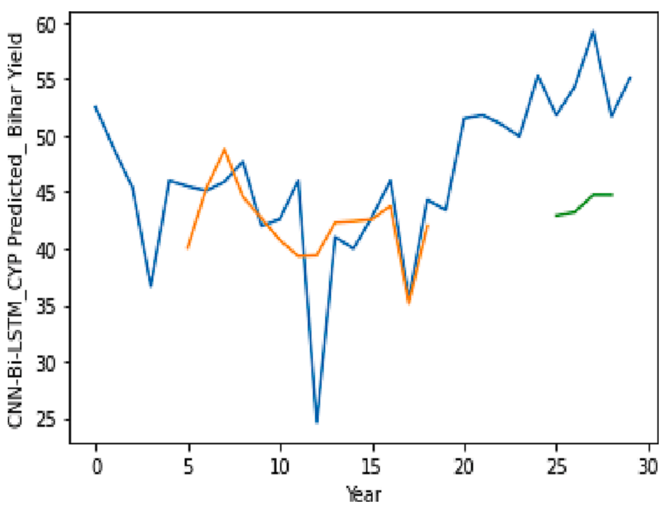


Fig. 5. Bihar Sugarcane Yield Prediction using CNN-Bi-LSTM_CYP Model.

based methods. The time sequence of 10 years was utilized to evaluate the forecasting and observe the increasing crop yield pattern for the major producing regions. In the experimentation, the proposed approach outperformed the existing methods with an RMSE value of 4.05 compared to other machine learning paradigms.

The limitation of the proposed work is the use of the sugarcane crop. The result might be different for another set of crop datasets. Another limitation exists in the case of feature selection that can also influence the network's training. In the future, this study can be extended by combining the hybrid technique of Deep Neural Network with XGBoost or optimization method to enhance the capability of the hybrid network for optimizing the selection of optimal hyperparameters to train the network.

Funding

This research receive no external funding.

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Preeti Saini: Conceptualization, Methodology, Software, Writing – original draft. **Bharti Nagpal:** Conceptualization, Methodology, Data curation, Supervision. **Puneet Garg:** Conceptualization, Visualization, Investigation. **Sachin Kumar:** Supervision, Validation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

The study is supported by the Russian Science Foundation Grant number: 23-21-10009.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.seta.2023.103263.

References

[1] Food and Agriculture Organization of UN. https://www.fao.org/home/en. [Accessed 5 January 2022].
[2] Food and Agriculture Organization of the United Nations. How to feed the world in 2050. https://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf. [Accessed 30 Jan 2022].
[3] FAO, IFAD, UNICEF, WFP, WHO. FAO, Rome. The State of Food Security and Nutrition in the World. Transforming Food Systems for Affordable Healthy Diets. 2020; 978-92-5-132901-6:320. 10.4060/ca9692en.
[4] Shao YE, Dai J-T. Integrated feature selection of ARIMA with computational intelligence approaches for food crop price prediction. Complexity 2018;2018: 1–17.
[5] Amrouk EM, Heckelei T. Forecasting international sugar prices: A Bayesian model average analysis. Sugar Tech 2020;22:552–62. https://doi.org/10.1007/s12355-020-00815-0.
[6] Vishwajith KP, Sahu PK, Dhekale B, Mishra P. Modelling and forecasting sugarcane and sugar production in India. Indian J Econ Dev 2016;12(1):71.
[7] Mishra P, Al Khatib MG, Sardar I, Mohammed J, Karakaya K, Dash A, et al. Modeling and Forecasting of Sugarcane Production in India. Sugar Tech 2021;23: 1317–24. https://doi.org/10.1007/s12355-021-01004-3.
[8] Ansarifar J, Wang L, Archontoulis SV. An interaction regression model for crop yield prediction. Sci Rep 2021;11:17754. https://doi.org/10.1038/s41598-021-97221-7.
[9] Oikonomidis A, Catal C, Kassahun A. Hybrid deep learning-based models for crop yield prediction. Appl Artif Intell 2022;36(1).
[10] Martínez-Ferrer L, Piles M, Camps-Valls G. Crop yield estimation and interpretability with Gaussian processes. IEEE Geosci Remote Sens Lett 2020;99: 1–5.
[11] Ministry of Statistics and Programme Implementation. https://www.mospi.gov.in/. [Accessed 5 January 2022].
[12] Department of Food & Public Distribution https://dfpd.gov.in/sugar.htm [Accessed 15 January 2022].

[13] Directorate of Economics and Statistics, Department of Agriculture, Cooperation and Farmers Welfare. https://aps.dac.gov.in/Home.aspx?ReturnUrl=%2f [Accessed 30 Jan 2022].
[14] Sanjeev, Verma U. ARIMA versus ARIMAX for sugarcane yield prediction in Haryana. Int J Agric Stat Sci 2016;12(2):327–34.
[15] Box GEP, Jenkins GM, Reinsel GC. Time series analysis; forecasting and control. 3rd Edition. Englewood Cliff, New Jersey: Prentice Hall; 1994.
[16] Bader EA. Economic modelling and forecasting of sugar production and consumption in Egypt. Int J Agric Econ 2017;2(4):96–109. https://doi.org/10.11648/j.ijae.20170204.12.
[17] Suman, Verma U. State space modelling and forecasting of sugarcane yield in Haryana, India. J Appl Nat Sci 2017; 9 (4):2036–2042. 10.31018/jans.v9i4.1485.
[18] Jeong S, Ko J, Yeom JM. Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea. Sci Total Environ 2022;802:149726. https://doi.org/10.1016/j.scitotenv.2021.149726.
[19] Nevavuori P, Narra N, Lipping T. Crop yield prediction with deep convolutional neural networks. Comput Electron Agr 2019;163(11):104859.
[20] Maldaner LF, Corrêdo LDP, Canata TF, Molin JP. Predicting the sugarcane yield in real-time by harvester engine parameters and machine learning approaches. Comput Electron Agr 2021;181:0168–1699. https://doi.org/10.1016/j.compag.2020.105945.
[21] Kumar S, Kumar V, Sharma RK. Sugarcane yield forecasting using artificial neural network models. Int J Artif Intell (IJAIA) 2015;6(5):51–68. https://doi.org/10.5121/ijaia.2015.6504.
[22] Kim N, Lee YW. Machine learning approaches to corn yield estimation using satellite images and climate data: A case of Iowa State. J Korean Soc Surv Geod Photogramm Cartogr 2016;34(4):383–90. https://doi.org/10.7848/ksgpc.2016.34.4.383.
[23] Tanut B, Waranusast R, Riyamongkol P. High accuracy pre-harvest sugarcane yield forecasting model utilizing drone image analysis, data mining, and reverse design method. Agriculture 2021;11(7):682. https://doi.org/10.3390/agriculture11070682.
[24] Clevers JGPW, Kooistra L, Van den Brande MMM. Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. Remote Sens 2017;9:405. https://doi.org/10.3390/rs9050405.
[25] Mohammad-Parsa H, Lu S, Kamaraj K, Slowikowski A, Haygreev CV. Deep learning architectures. In: Deep learning: concepts and architectures. Berlin/Heidelberg, Germany: Springer; 2019. p. 1–24.
[26] Bejo SK, Mustaffha S, Ishak W, Ismail WIBW. Application of artificial neural network in predicting crop yield: A review. J Food Sci Eng 2014;4:1–9.
[27] Patryk H, Magdalena P, Gniewko N. Selection of independent variables for crop yield prediction using artificial neural network models with remote sensing data. Land 2021;10:609. https://doi.org/10.3390/land10060609.
[28] Safa M, Samarasinghe S, Nejat M. Prediction of wheat production using artificial neural networks and investigating indirect factors affecting it: case study in Canterbury Province. New Zealand J Agr Sci Tech 2015;17(4):791–803.
[29] Haider SA, Naqvi SR, Akram T, Umar GA, Shahzad A, Sial MR, et al. LSTM neural network-based forecasting model for wheat production in Pakistan. Agron 2019;9:72. https://doi.org/10.3390/agronomy9020072.
[30] Suebsombut P, Sekhari A, Sureephong P, Belhi A, Bouras A. Field data forecasting using LSTM and Bi-LSTM approaches. Appl Sci 2021;11(24):11820. https://doi.org/10.3390/app112411820.
[31] Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput 2019;31:1235–70. https://doi.org/10.1162/neco_a_01199.
[32] Bali N, Singhla N. Deep learning based wheat crop yield prediction model in Punjab region of North India. Appl Artif Intell 2021;35:1304–28. https://doi.org/10.1080/08839514.2021.1976091.
[33] Timbadia DH, Sudhanvan S, Shah PJ, Agrawal S. Crop yield prediction for India using regression algorithms. Int Conf Adv Comput Data Sci 2021:241–51.
[34] Na-udom A, Rungrattanaubol J. A comparison of artificial neural network and regression model for predicting the rice production in lower Northern Thailand. In: Information Science and Applications. Springer; 2015. p. 745–52. https://doi.org/10.1007/978-3-662-46578-3_88.
[35] Gopal PSM, Bhargavi R. A novel approach for efficient crop yield prediction. Comput Electron Agr 2019;165(2):104968. https://doi.org/10.1016/j.compag.2019.104968.
[36] Shastry KA, Sanjay HA. Hybrid prediction strategy to predict agricultural information. Appl Soft Comput 2020;98(3):106811. https://doi.org/10.1016/j.asoc.2020.106811.
[37] Maya GPS, Bhargavi R. Performance evaluation of best feature subsets for crop yield prediction using machine learning algorithms. Appl Artif Intell 2019;33:621–42. https://doi.org/10.1080/08839514.2019.1592343.
[38] Meng X, Liu M, Wu Q. Prediction of rice yield via stacked LSTM. Int J Agric Environ Inf Syst (IJAEIS) 2020; 11(1):86-95. http://doi.org/10.4018/IJAEIS.2020010105.
[39] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Computer Science. CoRR.2014; arXiv:1409.1556.
[40] Rhanoui M, Mikram M, Yousfi S, Barzali S. A CNN-BiLSTM model for document-level sentiment analysis. Mach Learn Knowl Extr 2019;1(3):832–47. https://doi.org/10.3390/make1030048.
[41] Ketu S, Mishra PK. India perspective: CNN-LSTM hybrid deep learning model-based COVID-19 prediction and current status of medical resource availability. Soft Comput 2022;26:645–64. https://doi.org/10.1007/s00500-021-06490-x.
[42] Huang C-J, Kuo P-H. A deep CNN-LSTM model for particulate matter (PM2. 5) forecasting in smart cities. Sensors 2018;18(7):2220. https://doi.org/10.3390/s18072220.
[43] Bhojani SH, Bhatt N. Wheat crop yield prediction using new activation functions in a neural network. Neural Comput Appl 2020;32:13941–51. https://doi.org/10.1007/s00521-020-04797-8.
[44] IBM Cloud Education. Recurrent Neural Networks. https://www.ibm.com/cloud/learn/recurrent-neural-networks [Accessed 30 Jan 2022].
[45] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80. https://www.bioinf.jku.at/publications/older/2604.pdf.
[46] Fei H, Tan F. Bidirectional grid long short-term memory (BiGridLSTM): A method to address context-sensitivity and vanishing gradient. Algorithms 2018;11(72):1. https://doi.org/10.3390/a11110172.
[47] Hanane Elfaik, El Habib Nfaoui. Deep bidirectional LSTM network learning-based sentiment analysis for Arabic text. J Intell Syst 2021; 30:395-412. 10.1515/jisys-2020-0021.
[48] Priyadarshini I, Cotton C. A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis. J Supercomput 2021;77:13911–32. https://doi.org/10.1007/s11227-021-03838-w.
[49] Chen B, Zheng H, Wang L, Hellwich O, Chen C, Yang L, et al. A joint learning Im-BiLSTM model for incomplete time-series Sentinel-2A data imputation and crop classification. Int J Appl Earth Obs Geoinf 2022;108:1569–8432. https://doi.org/10.1016/j.jag.2022.102762.
[50] Singh M, Tyagi V, Gupta PK, Flusser J, Ören T, Sonawane VR. Advances in Computing and Data Sciences. ICACDS 2021. Commun Comput Inf Sci 2021; 1441. 10.1007/978-3-030-88244-0_23.
[51] Mishra P, Yonar A, Yonar H, Kumari B, Abatoble A, SankarDas S, et al. State of the art in total pulse production in major states of India using ARIMA techniques. Curr Res Food Sci 2021;4:800–6. https://doi.org/10.1016/j.crfs.2021.10.009.
[52] Khaki S, Wang L, Archontoulis SV. A CNN-RNN framework for crop yield prediction. Front Plant Sci 2020;11:1750. https://doi.org/10.3389/fpls.2019.01750.