# Journal of Agrometeorology

ISSN : 0972-1665 (print), 2583-2980 (online)
Vol No. 24(2) : 179-185 (June 2022)

https://journal.agrimetassociation.org/index.php/jam


*Research Paper*

# A comprehensive approach on predicting the crop yield using hybrid machine learning algorithms


**KRITHIKHA SANJU SARAVANAN and VELAMMAL BHAGAVATHIAPPAN**

*Department of Computer Science and Engineering, College of Engineering Guindy, Anna University, Chennai, India*
*Corresponding Author E-mail id: krithikhasanju3008@gmail.com



## ABSTRACT

Crop yield prediction is a complex task which uses historical data to predict how much yield can be obtained in a particular year. To predict accurate crop yield, a novel deep neural network named crop yield predicting deep neural network with XGBoost regression and AdaBoost regression algorithms were used. Further, in this research work, prediction models proposed are hybrid models, namely PCA-XGBoost, PCA-AdaBoost and the LSTM based Stacked Auto Encoder – Crop Yield Predicting Deep Neural Network (LSAE-CYPDNN) model for predicting the crop yield. The error metrics like Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) were evaluated for the hybrid models. The result shows that the proposed hybrid LSAE-CYPDNN model yields much less MAE, MAPE and RMSE compared to the models PCA-AdaBoost and PCA-XGBoost.



**Keywords:** *Principal component analysis, XGBoost regression, AdaBoost regression, Stacked Auto Encoder, Deep Neural Network.*


Food demand is expected to increase between 59% and 98% by 2050 (Valin *et al.*, 2014). Large amounts of crops must be grown to meet the increasing demand for food. To attain the goal, it is necessary to increase the crop yield with low field farmland. The yield prediction can therefore help the planning authorities with essential data in order to make sensible and correct decisions. In the process of agricultural planning, agro-economists have to provide simple and accurate calculation methods to forecast yields. Crop yield prediction depends on input aspects like location, irrigation techniques, temperature, etc. and mainly the crop yield forecasting algorithms. It is possible to obtain high precision of crop yield by implementing with suitable inputs and models without disrupting the essence and structures of agricultural production.

Agricultural researchers were examining stronger yield forecasting which depends on many factors. However, no baseline data collection is accessible for agricultural research as well as ranges for different locations, type of crop, climate and irrigation technique. Researchers use data-driven systems to obtain accurate prediction using the available data. Machine Learning (ML) algorithms play a key role in achieving greater precision in the data-driven model. Though there are major advances in ML and implementation in several areas, ML methods have some constraints when utilized in a solely data-driven manner. The accuracy of the forecasts and the accommodations developed by ML algorithms depends on the level of the data, the representative of the system and the dependence of the input as well as target variables in the datasets gathered (Khaki & Wang, 2019).

Recently, there has been an increase in the use of cloud computing and big data for agricultural applications since any health issues in plants can be analysed by experts over the internet by using the symptoms (Ali et al., 2019) which also affects the crop yield. Some researchers utilized crop yield prediction regression models to evaluate their usefulness with other models. For better crop yield models, analyse the classical statistical models with other models (Barbedo, 2018). In this research, high computational regression based machine learning algorithms and a novel deep neural network are utilized in combination for predicting the crop yield.

In the related work, the analysis of few crop yield prediction research works using various predictive modelling techniques was explored. Khaki and Wang (2019) have evaluated a novel deep learning algorithm for Syngenta Crop challenge. Similarly, Chlingaryan *et al.,* (2018) have also addressed scientific advancement in the last 15 years on machine-based learning


**Article info - DOI:** https://doi.org/10.54386/jam.v24i2.1561
*Received: 1 March 2022; Accepted: 29 April 2022; Published online: 28 May 2022*

techniques for precise crop yield prediction based on the assessment of nitrogen status. As a limitation, Chlingaryan *et al.*,(2018) found that, the sensor system has less framework and used only simple machine learning techniques. A simulation based model has been proposed by Mohanty *et al.*, (2015) for calculating the wheat productivity based on the atmospheric temperature and carbon di oxide levels. Evaluated the research in Bhopal region of Madhya Pradesh and found that there was a positive relationship between carbon di oxide and grain yield. Also found that there was a negative relationship between atmospheric temperature and grain yield. Hence this study has proved that there is a huge impact on the wheat yield due to environmental factors. Whereas, the Maya and Bhargavi,(2019) have used the back propagation learning algorithm with Feed Forward Artificial Neural Network to predict the exact yield of paddy crops. Further, Crane-Droesch (2018), Khaki and Wang (2019) used deep neural network for crop yield prediction. Khaki and Wang (2019) suggested that error rate for crop yield prediction should be reduced and suggested for new enhanced technique. From the above existing research works, it is seen that crop yield prediction efficiency and accuracy is not adequate so new hybrid model is required for better accuracy and precision.

## MATERIALS AND METHODS

The objective of the research is to predict crop yields with high precision. Regression and DNN based machine learning algorithms and their performance metrics are used to determine the effectiveness of the proposed hybrid models. The system consists of the following major phases like dataset selection, then data pre-processing like data imputation, solving positive skewed problems, one-hot encoding. Followed by dimensionality reduction and feature extraction, then training and testing. Finally, the error metrics are used to evaluate the machine learning and deep learning techniques. The research work flow is divided into three sections and is given below,

1) Crop yield dataset →Data pre-processing (Data imputation, Solving positive skewed problem, One hot encoding)→Feature extraction (Principal Component Analysis)→XGBoost regression based prediction model and evaluating the model using evaluation metrics.

2) Crop yield dataset→Data pre-processing (Data imputation, Solving positive skewed problem, One hot encoding)→Feature extraction (Principal Component Analysis)→AdaBoost regression based prediction model and evaluating the model using evaluation metrics.

3) Crop yield dataset→Data pre-processing (Data imputation, Solving positive skewed problem)→Feature extraction (LSTM based Stacked Auto Encoder)→Deep learning based prediction model (DNN) and evaluating the model using evaluation metrics.

### *Dataset description*

In this work, a dataset that consists of historical data of crop yield from all over Indian states and all districts of India were used between the year 1995 to 2015. The dataset consists of seven instances (data columns - state_name, district_name, crop_year, season, crop_name, area and production) with 246091 different Attributes (rows). The research work considered nearly 124 types of crops that has been grown in all over India. The basic features of the samples are presented in Table 1. The source of the dataset used in this research work is https://www.kaggle.com/code/anjali21/indian-production-analysis-and-prediction/data.

### *Data imputation*

The first step of pre-processing is to check for the missing value in the dataset. The missing data in the dataset causes problems such as biased or skewed, complicated data management and interpretation with performance reductions. So, mean imputation has been used in this work. This patches the missing data in the dataset and updates the missing data by the substituted data.

### *Positive skewed problems*

Once data imputation has been performed using the mean mathematical technique, data is now distributed on one side. As a result, the long tail is on the positive side of the peak and skewed to the right. So, it's called a positive skewed problem. Hence, this target variable is totally skewed on one side of the distribution, which would give a bias in the result, so it is necessary to convert the target variable into a normal distribution. In order to make a normal distribution, mathematical log operations are used and made the target variable into a normal distribution format.

### *One hot encoding*

The third pre-processing technique is one hot encoding technique which is used for obtaining better accuracy. Machine learning algorithms cannot directly work with categorical data, so categorical data is converted to numeric data by one-hot encoding. A one-hot encoder represents categorical variables as binary vectors. This encoder helps to map the categorical values to integer values. As a binary vector, each integer value is represented. Using the one-hot encoding technique all the features in the dataset is converted to the numerical data. Further, using this technique the seven-feature column is extended to the 800-feature column.

### *Feature extraction*

Once all pre-processing is done, the data needs to be scaled down because the column of seven features is expanded to the column of 800 features. So, there is a certain need for dimensionality reduction. The feature extraction is also known as dimension reduction, since it reduces the dimensionality of data by selecting important features in data. While the initial information set is still accurate and original. Here for the dimension reduction, the Principal Component Analysis (PCA) is used. PCA is a statistical method which uses an orthogonal transformation to transform a set of observed possible associated variables. The 800 attributes from the dataset are reduced to 150 principal features using the PCA technique. Since the PCA components are orthogonal to each other, they are not correlated and they are independent. The pseudocode for the PCA technique used is given below.

1. Input: a D- dimensional training set X = X = $\{x_1, x_2, \ldots, x_n\}$ , N =

Dataset Size.

2. Compute the mean $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$

3. Compute the covariance matrix Cov $(x) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x}_i)(x_i - \bar{x}_i)^T$

4. From computed Covcompute the eigenvectors $\xi_1, \xi_2, \ldots, \xi_D$ and their corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_D$ and then they are sorted, such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_D \geq 0$

5. From computed Eigen vectors and Eigen values, select value with more variance for the dataset. Using those Eigen value and Eigen vector, compute new lower dimensional data.

6. For any $x \in \mathbb{R}^D$ its new lower dimensional representation is:

$$y = \left(\xi_1^T(x-\bar{x}),\ \xi_2^T(x-\overline{x}), \ldots, \xi_d^T(x-\bar{x})\right)^T \in \mathbb{R}^d,$$

7. The original can be approximated as

$$x \approx \bar{x} + (\xi_1^T(x-\bar{x}))\xi_1 + (\xi_2^T(x-\bar{x}))\xi_2 + \cdots + \xi_d^T(x-\bar{x}))\xi_d.$$

***Regression***

Regression analysis is the evaluation of how one or more predictors depend on a response variable, like how crop yield alterations vary as inputs like a quantity of irrigation or type of seed. In this research the two models used for regression are XGBoost regression and AdaBoost. The result obtained from PCA is given as the input for these models. Once the data is trained, remaining data is sent to the testing phase which also uses the same machine learning based regression algorithms for testing. The splitting ratio of a dataset for training is 80%, and testing is 20%. The regression uses a dependent variable and an independent variable to predict future crop yields.

***XGBoost regression***

XGBoost is a supervised learning based boosting algorithm and an ensemble algorithm dependent on gradient boosted trees (Mo et al., 2019). By introducing a regularization term, it will eliminate over-fitting. The $x_i$ is determined by Gunay et al.(2013).

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i)$$

Where $\hat{y}_i^{(t)}$ - the final tree model, $\hat{y}_i^{(t-1)}$- earlier generated tree model, $f_k(x_i)$ - recently generated tree model, t -total number of base tree models.

The depth, as well as trees numbers, are significant parameters for the XGBoost algorithm. The XGBoost algorithm will accommodate any variables as inputs, but the output variable must be discrete, like binary variables (Mo et al., 2019). A new tree is formed along the direction of the loss function's negative gradient. The loss becomes smaller and smaller when the number of tree models increases. The features extracted from PCA are given as input and the XGBoost based regression output for predicting the crop yield have been obtained successfully.

***AdaBoost regression***

AdaBoost is also known as Adaptive Boosting (Xiao *et al.*, 2019). AdaBoost is actually intended for challenges with regression and classification. An AdaBoost regressor is a meta-estimator which starts by installing a regressor on the initial dataset and then suits subsequent regressor clones on the similar dataset but changes the amounts of instances as per the actual prediction error. Further, it combines multiple weak regressors to increase the accuracy of the main model (Wu et al., 2020). Here also the output from PCA is given as the input and the AdaBoost based regression output for predicting the crop yield have been obtained successfully.

***Deep learning***

A deep neural network (DNN) is a multi-layered artificial neural network (ANN). Where the multiple layers will be between the input and output layers (Hinton *et al.,* 2012). Whereas, the basic idea of driving ANN is the design, which imitates how the brain functions as neuronal interconnections (Rumelhart & Hinton, 1986). The DNN is good in object localization (Girshick *et al.,* 2014), visual classification (Wu *et al.*, 2020; A. Krizhevsky, I. Sutskever, 2012) and speech recognition (Längkvist *et al.*, 2014). In this research, a DNN was utilized to establish associations between input and output variables as a multivariate regression model. DNNs training involves fine-tuning as well as pre-training (Merkel *et al.*, 2018). The regression evaluation function in DNN is used to evaluate the accuracy of the crop yielding prediction model. The LSTM based Stacked Auto Encoder (LSAE) – Deep Neural Network learning model is proposed and this enhanced feature extraction with deep learning model known as LSAE-CYPDNN. Stacked auto encoder is used for automatic unsupervised feature extraction. Hence LSAE extracts features automatically and the extracted features are given as the input to the DNN for predicting the crop yield. 1024 artificial neurons are used. Such artificial neurons take one or more inputs, which are combined and added together by values called "weights". This value is then transferred to a non-linear function, identified as an activation function, which will become the output of the neuron. To train this DNN model, total 24 hidden layers were functioned which gives best model for predicting the crop yield. The proposed DNN uses multiple epochs (Reddy & Rao, 2020, François, 2020). The number of epochs in this model is 20.

Adam is an optimization algorithm that could be utilized to modify network weights iteratively dependent on training data rather than the traditional stochastic gradient descent technique (Kouzehgar *et al.*, 2019). Where, Adam (Kingma & Ba, 2014) is an adaptive learning rate optimization algorithm used in the proposed model.

An activation function is a non-linear function that a neuron applies to introduce non-linear properties into the network. A non-linear relation indicates a transition in the first variable will not automatically equate to a constant change in the second. The activation function in a neural network is liable for translating the cumulative weighted input from the node into the activation node or providing output from the input. ReLU is known as Rectified Linear Unit and it a most used activation function in data science (Tang *et al.*, 2018). ReLU non-linear, monotonic and convex function used

**Table 1:** Crop yield prediction dataset sample

| State_Name | District_Name | Crop_Year | Season | Crop | Area | Production |
|---|---|---|---|---|---|---|
| Andaman and Nicobar Islands | NICOBARS | 2000 | Kharif | Arecanut | 1254 | 2000 |
| Andhra Pradesh | ANANTAPUR | 1997 | Rabi | Ragi | 600 | 1000 |
| Bihar | GOPALGANJ | 2006 | Rabi | Maize | 5626 | 11134 |
| Chandigarh | CHANDIGARH | 2008 | Rabi | Wheat | 600 | 2700 |
| Haryana | KAITHAL | 2012 | Whole Year | Sugarcane | 3511 | 349000 |
| Karnataka | BELGAUM | 2013 | Kharif | Bajra | 12444 | 8204 |
| Tamil Nadu | VILLUPURAM | 2013 | Whole Year | Turmeric | 2252 | 5040 |
| West Bengal | PURULIA | 2014 | Rabi | Urad | 220 | 113 |

**Table 2:** Future crop yield prediction results sample of CYPDNN model

| State_Name | District_Name | Crop_Year | Season | Crop | Area | Production |
|---|---|---|---|---|---|---|
| Tamil Nadu | THANJAVUR | 2022 | Kharif | Rice | 178011 | 684316 |
| Tamil Nadu | CUDDALORE | 2023 | Kharif | Rice | 125207 | 540382 |
| Tamil Nadu | PERAMBALUR | 2022 | Whole Year | Cotton(lint) | 14000 | 34431 |
| Tamil Nadu | MADURAI | 2023 | Whole Year | Cotton(lint) | 10400 | 24753 |
| Tamil Nadu | SALEM | 2022 | Whole Year | Cotton(lint) | 12500 | 37248 |
| Tamil Nadu | DHARMAPURI | 2022 | Whole Year | Sugarcane | 2424 | 273610 |
| Tamil Nadu | TIRUVANNAMALAI | 2022 | Whole Year | Sugarcane | 19500 | 2340000 |
| Tamil Nadu | TIRUCHIRAPPALLI | 2022 | Rabi | Groundnut | 5100 | 15390 |
| Tamil Nadu | NAMAKKAL | 2023 | Rabi | Groundnut | 5000 | 17480 |

**Table 3:** Comparison for Mean absolute error (MAE), Root mean squared error (RMSE) and Mean absolute percentage error (MAPE)

| Proposed models | Mean absolute error | Root mean squared error | Mean absolute percentage error |
|---|---|---|---|
| XGBoost | 11.3% | 13% | 5% |
| AdaBoost | 12.6% | 15% | 10% |
| CYPDNN | 4.2% | 6.5% | 3% |

**Table 4:** Comparison of proposed deep learning model with existing models using MAPE error metric

| Research Work | MAPE |
|---|---|
| You et al.,(2017) | 3.41% |
| Oliveira et al., (2018) | 9.8% |
| Laxmi and Kumar,(2011) | 4.4% |
| Proposed Method | 3% |

to catch more complex patterns on the positive domain and obtain high values with less error. The ReLU activation function is affected by dying ReLU problem. Leaky ReLU activation function is used to overcome the dying ReLU problem. The parametric ReLU is used as activation function to solve the issues created by the hyperparameter of Leaky ReLU activation function. Further Gaussian Error Linear units (GeLU) is non-linear, non-monotonic and non-convex function used to catch more complex patterns at all points and is used in the proposed model which shows better accuracy than others activation functions.

***Loss calculation***

In this research, three techniques such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) are used to find the error percentages of the predicting models.

***Mean Absolute Error (MAE)***

A mean absolute error metric is used to calculate the accuracy of the continuous variable. The MAE is used to calculate the similarity of the forecast for the possible results. Sum of all absolute errors is the Mean Absolute Error (MAE). The absolute error is the amount of error during observation. It is the difference between the predicted values and actual values. The equation for the Mean Absolute Error (MAE) is represented as,

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|p_i - a_i|$$

Where, n = the number of observed errors, Σ = summation (which means "sum of all"), $a_i$= actual value, $p_i$ = predicted value, $|p_i - a_i|$ = absolute errors.

***Mean absolute percentage error (MAPE)***

MAPE is an evaluation measure that is used to measure how accurate the prediction system is. It measures the prediction accuracy as a percentage. The equation for the Mean Absolute Percentage Error (MAPE) is represented as,

**Table 5:** Comparison of hyperparameters

| Activation function | RMSE (%) | Optimizer | RMSE (%) |
|---|---|---|---|
| Tanh | 11.15 | SGD | 10.89 |
| Sigmoid | 9.76 | RMSPROP | 10.25 |
| Relu | 7.64 | ADAGRAD | 8.45 |
| Leaky Relu | 7.43 | ADAM | 6.50 |
| Parametric Relu | 7.18 | | |
| Gelu | 6.50 | | |

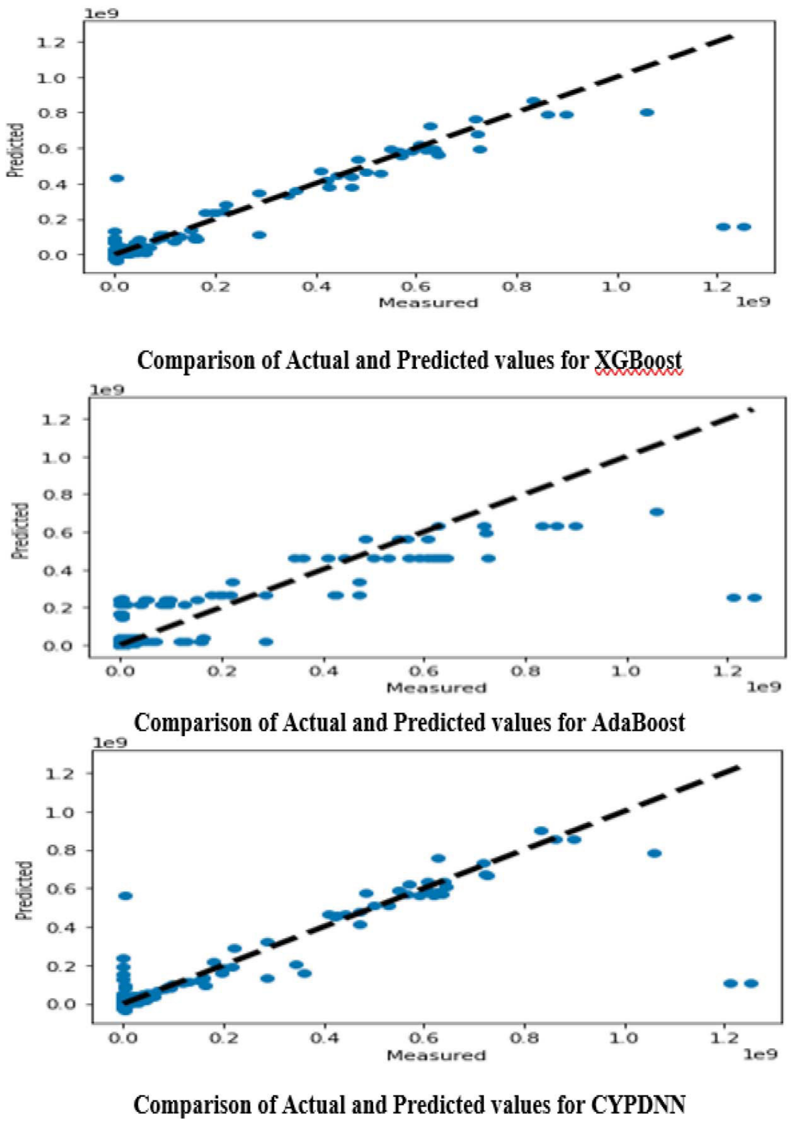


**Fig. 1:** Comparison of actual and predicted values for all the three models

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\frac{|a_i - p_i|}{|a_i|}$$

***Root mean squared error (RMSE)***

The standard deviation of prediction errors (residuals) is the Root Mean Square Error (RMSE) (Sinha et al., 2021). RMSE calculates the residuals distributions. It shows the best fit line of the concentrated data. To find the accuracy, the root mean square is taken for the error, where the error occurred between the observed value and predicted value. The formula for the RMSE is given below.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(p_i - a_i)^2}{n}}$$

## RESULTS AND DISCUSSION

Once crop yield dataset is trained with 80% data and tested with 20% data, performance evaluation is performed using the error metrics. Also, future crop yield predictions were done for the best CYPDNN model and the sample is shown in Table 2. The error metrics are evaluated for all the three proposed crop yield predicting hybrid models.

### *Comparison of XGBoost Regression, AdaBoost Regression and CYPDNN using MAE, RMSE and MAPE*

The error metrics are compared for all the three models and the proposed CYPDNN model has the least error for all the error metrics while the highest error rates are recorded for the AdaBoost model. The error metrics comparisons are shown clearly in Table 3.

The comparison of the existing techniques with the proposed methodology with respect to MAPE is shown in Table 4. While the research works have different MAPE for different crops, the average is given. The proposed algorithm has very less error when compared with existing algorithms.

### *Comparison of Hyperparameters*

The comparison is also performed to check the difference between the activation functions. This work uses GeLU as the activation function. Other activation functions like ReLU, Tanh and Sigmoid are also implemented and compared with respect to RMSE is seen in Table 5. It is seen that the GeLU works effectively for the selected model and dataset and has less error rate. This is the reason that this particular activation function is implemented. This work also uses ADAM optimizer, while the other optimizers compared are SGD, ADAGRAD, and RMSPROP. It can be seen that ADAM is more efficient by a large margin and it is also shown in Table 5.

Performance is evaluated for the proposed model, where the actual and predicted value is examined. The proposed LSAE-CYPDNN provides better prediction when compared with the other two models AdaBoost and XGBoost. The actual value and prediction value for the crop yield prediction for XGBoost, AdaBoost and the proposed LSAE-CYPDNN algorithm is represented in Fig 1.

## CONCLUSION

In this research work, three hybrid models were proposed namely PCA-XGBoost, PCA-AdaBoost and LSAE-CYPDNN model to predict the accurate crop yield. The proposed LSAE-CYPDNN algorithm provides the most minimum optimal error and helps to increase the accurate prediction. The predicted error is calculated using the error metrics. Three errors metric is used namely Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE). The results were compared with all other proposed hybrid models and with existing algorithms. The proposed Deep learning algorithm is better and has less error when compared to existing techniques. Finally, the result shows that the proposed hybrid LSAE-CYPDNN model gives minimum optimal error when compared to the other two proposed hybrid models PCA-AdaBoost and PCA-XGBoost.

***Conflict of Interest Statement:*** The authors declare that there is no conflict of interest.



## REFERENCES

Ali, M. M., Bachik, N. A., Muhadi, N. A., Yusof, T. N. T., & Gomes, C. (2019). Non-destructive techniques of detecting plant diseases: A review. *Physiol. Mol. Plant Pathol.*, 108: 101426.

Barbedo, J. G. A. (2018). Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput. Electron. Agric.*, 153: 46-53.

Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.*, 151: 61-69.

Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environ. Res. Lett.*, 13(11): 114003.

François, O. (2020). A multi-epoch model for the number of species within genera. *Theor. Popul. Biol.*, 133: 97-103.

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *In*: Proceedings of the IEEE conference on computer vision and pattern recognition, p.580-587.

Gunay, H. B., O'Brien, W., & Beausoleil-Morrison, I. (2013). A critical review of observation studies, modeling, and simulation of adaptive occupant behaviors in offices. *Build. Environ.*, 70: 31-47.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., ... & Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal processing magazine, 29(6), p.82-97.

Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. *Front. Plant Sci.*, 10: 621.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *Arxiv Prep.*, 1412: 6980.

Kouzehgar, M., Tamilselvam, Y. K., Heredia, M. V., & Elara, M. R. (2019). Self-reconfigurable façade-cleaning robot equipped with deep-learning-based crack detection based on convolutional neural networks. *Autom. Constr.*, 108: 102959.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Adv. *Neural Inf. Process. Syst.*, 25: 1097-1105.

Längkvist, M., Karlsson, L., & Loutfi, A. (2014). A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognit. Lett.*, 42: 11-24.

Laxmi, R. R., & Kumar, A. (2011). Weather based forecasting model for crops yield using neural network approach. *Stat. Appl.*, 9(1&2): 55-69.

Li, Y., Guan, K., Yu, A., Peng, B., Zhao, L., Li, B., & Peng, J. (2019). Toward building a transparent statistical model for improving crop yield prediction: Modeling rainfed corn in the US. *Field Crops Res.*, 234: 55-65.

Gopal, P. M., & Bhargavi, R. (2019). A novel approach for efficient crop yield prediction. *Comput. Electron. Agric.*, 165: 104968.

Merkel, G. D., Povinelli, R. J., & Brown, R. H. (2018). Short-term load forecasting of natural gas with deep neural network regression. *Energies*, 11(8): 2008.

Mo, H., Sun, H., Liu, J., & Wei, S. (2019). Developing window behavior models for residential buildings using XGBoost algorithm. *Energy Build.*, 205: 109564.

Mohanty, M., Sinha, N. K., Hati, K. M., Reddy, K. S., & Chaudhary, R. S. (2015). Elevated temperature and carbon dioxide concentration effects on wheat productivity in Madhya Pradesh: a simulation study. *J. Agrometeorol.*, 17(2): 185.

Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Comput. Electron. Agric.*, 163: 104859.

Oliveira, I., Cunha, R. L., Silva, B., & Netto, M. A. (2018). A scalable machine learning system for pre-season agriculture yield forecast. *Arxiv Prep.*, 1806: 09244.

Reddy, M. K., & Rao, K. S. (2020). Excitation modelling using epoch features for statistical parametric speech synthesis. *Comput. Speech Lang.*, 60: 101029.

Rumelhart, D. E., Hinton, G. E., & Williams, R. (1986). J.(1986). Learning internal representations by error propagation. *Parallel Distrib. Proces.*, 1.

Sinha, N. K., Mohanty, M., Somasundaram, J., Chaudhary, R. S., Patra, H., Hati, K. M., ... & Prabhakar, M. (2021). Maize productivity analysis in response to climate change under different nitrogen management strategies. *J. Agrometeorol.*, 23(3): 279-285.

Tang, Z., Luo, L., Peng, H., & Li, S. (2018). A joint residual network with paired ReLUs activation for image super-resolution. *Neurocomputing*, 273: 37-46.

Valin, H., Sands, R. D., Van der Mensbrugghe, D., Nelson, G. C., Ahammad, H., Blanc, E., ... & Willenbockel, D. (2014). The future of food demand: understanding differences in global economic models. *Agric. Econ.*, 45(1): 51-67.

Wu, Y., Ke, Y., Chen, Z., Liang, S., Zhao, H., & Hong, H. (2020). Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. *Catena*, 187: 104396.

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., & Chen, Z. (2019). Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sens. Environ.*, 233: 111358.

You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep gaussian process for crop yield prediction based on remote sensing data. *In*: Thirty-First AAAI conference on artificial intelligence.