Indian Journal of Ecology (2022) 49(3) (SI): 1005-1010
DOI: https://doi.org/10.55362/IJE/2022/3628

Manuscript Number: 3628
NAAS Rating: 5.79

Crop Production Prediction Models in Indian Agriculture: Possibilities and Challenges

Rohini Bhat, Ritika Joshi and Vedpriya Arya[1]

Patanjali Herbal Research Department, Patanjali Research Institute, Haridwar-249 405, India
[1]Department of Allied Sciences, University of Patanjali, Patanjali Yogpeeth, Haridwar-249 405, India
E-mail: rohini.bhat@prft.in

Abstract**:** Agriculture sector is an important contributor to Indian economy. Wheat yield has decreased from 3.53 kg ha^{-1} (FY 2019) to 3.42 kg ha^{-1} (FY 2020). Crop production is affected by a lot of factors including climatological factors (temperature, precipitation), soil type and seed quality. Crop production prediction is an important aspect for farmers, agro based industries and policy makers. Many technological tools based on data mining models are being developed to draw correlation agricultural datasets and crop yield. The prediction methodology involves learning the pattern of crop yield during a set of conditions based on the previous years' data and thus predicting the yield in the current set of environmental parameters. The present paper focuses on commonly used models for crop yield prediction in Indian agriculture including artificial neural network, bayesian belief network, support vector machine, decision tree regression analysis, random forest, least absolute shrinkage regression operator and elastic net. A comparison of their efficiency has been drawn based on the previous studies.

Keywords: Agriculture sector, Climatological factors, Crop yield, Data mining, Prediction techniques

The agriculture sector is the backbone of India's food supply and economy. Many factors such as temperature, radiation, water availability and other environmental conditions affect the outcome of agriculture by influencing crop growth and development (Balkrishna et al 2021). Crop yield is important to farmers and other agriculture-related organizations and is critical to ensure the food security of a growing population, livelihoods of farmers and other agriculture-based organizations. India has faced several natural disasters such as drought and flooding resulting in loss of agricultural produce. Predicting agricultural yield ahead of time can play important role for quantitative and financial assessment of crop at the field level, planning for timely storage, transportation and marketing of agricultural products, fixing of sale prices by farmers and minimum support prices by policy makers and logistics planning by the agriculture-based industries (Jaiswal and Agrawal 2020). The progress in the latest technologies for global satellite observations and statistical models has led to the development of tools for crop prediction. Crop yield is predicted using various machine learning techniques based on mathematical and statistical methods. These tools help in making the relationship between meteorological factors and crop yield (Josephine et al 2020, Reddy and Kumar 2021, Vishal et al 2021). Data mining is an important approach involving database analysis, statistical analysis, and pattern recognition and thus plays a crucial role in extracting important information from large data. The data used for prediction includes meteorological data (temperature, rainfall, relative humidity), remote sensing data, and field statistics. Different data mining tools used for crop prediction include Artificial Neural Networks (ANN), Support Vector Machines (SVM) (Balkrishna et al 2020). A huge number of databases are used as input for crop prediction. Thus, choice of data mining tool, which can efficiently process the data, is critical for the accuracy of yield prediction. The present study focuses to discuss different data mining methods used for crop yield prediction and comparison of their accuracy.

Process of crop yield prediction: In general crop yield prediction methodology contains two phases- training phase (data is collected and pre-processed) and test phase (rules are generated which result in the prediction of yield as output) (Lata and Chaudary 2019). The crop yield prediction process is highlighted in Figure 1.

Collection of dataset: The first step in crop prediction is to feed the raw information. Crop yield is affected by agro-climatic factors which are further dependent on the field under study and agricultural practices. In India, climatic information of different places is collected by Indian Meteorological Department. The data is accessible without any charge.

Database pre-processing: The collection of data is followed by pre-processing the dataset for junk or noise removal. This includes arranging the data in a single file. The standard pattern includes columns with information of year, area, production, yield (production/area), monthly mean temperature, monthly mean rainfall, soil moisture, pH, etc. The data file is usually saved in .csv format.

Classification of data: The third step is to divide the data information into training and testing datasets. The splitting of data is done by train test split (scikit-learn). The training dataset contains useful information for training datasets to predict yield while the testing dataset employs the rest of the information for analyzing the performance of the system.

Crop yield prediction models used in India: Data mining tools used in agriculture are mainly of predictive type. A lot of data mining tools are currently being used for better interpretation, comprehension and making successful prediction based on large input data.The most common data mining tools used particularly in crop yield prediction in India are artificial neural network, bayesian belief network, support vector machine, decision tree, regression analysis, least absolute shrinkage regression operator (LASSO), elastic net (ENET) and random forest.

Artificial neural network: An artificial neural network (ANN) is an information processing model based on the working of the brain where a network of interconnected neurons processes the information (Fente and Singh 2018). ANN comprises a large number of simple processing units connected to form a complex network. Each of the units is capable of receiving and sending signals efficiently and has a standard topology in which one layer receives inputs, the other generates outputs and there are a lot of hidden layers in between. ANN works on supervised learning and is trained once to predict future patterns based on the pattern recognition of previous data used as input (Mahto et al 2021). This is the same as the human brain learning by examples. The accuracy of ANN increases with the amount of raw data provided. The mathematical model of ANN includes information about inputs, processing functions, and outputs.ANN is proving to be an important tool for crop production prediction. Using yield history and the responsible factors as input, the model can be configured for automatic learning of the pattern. The accuracy of ANN increases with the amount of raw data provided. The mathematical model of ANN includes information about inputs, processing functions, and outputs. ANN is proving to be an important tool for crop production prediction. Using yield history and the responsible factors as input, the model can be configured for automatic learning of the pattern. ANN efficiently resolves the nonlinear relation complexity existing between crop production and various predictor parameters provided as input. This can be used to generate predictions of the flood, pest attack, and rainfall patterns. Rice yield in Maharashtra was predicted using ANN with the accuracy of 97.5% and specificity of 98.1 (Gandhi et al 2016a).

Regression analysis: Regression modelling is based on evaluating the dependence of a dependent variable (predictant) on one or more independent variables (predictors) (Gonzalez-Sanchez et al 2014 Kumari et al 2016, Champaneri et al 2020). In crop yield prediction, yield is the dependent variable and all other agro climatic factors like temperature, rainfall, relative humidity, soil characteristics are considered as independent variables. Multiple regressions are employed when more than one independent variable is used as input. Stepwise multiple linear regressions (SMLR) is a regression technique where potential variables are added and weakly correlated variables are removed at each step based on the p-value and checked for statistical significance. Kadbhane and Manekar (2020) developed Agro-Climatic Grape Yield (ACGY) model based on multi regression analysis. Climatic factors for every month were used as input parameters. The performance of the statistical model was tested in the terms of discrepancy ratio (1.03), the standard deviation of discrepancy ratio (0.19), mean percentage error (0.03%), and standard deviation of mean percentage error (0.19). The model showed an increase in grape yield in the year 2050 in comparison to the current yield.

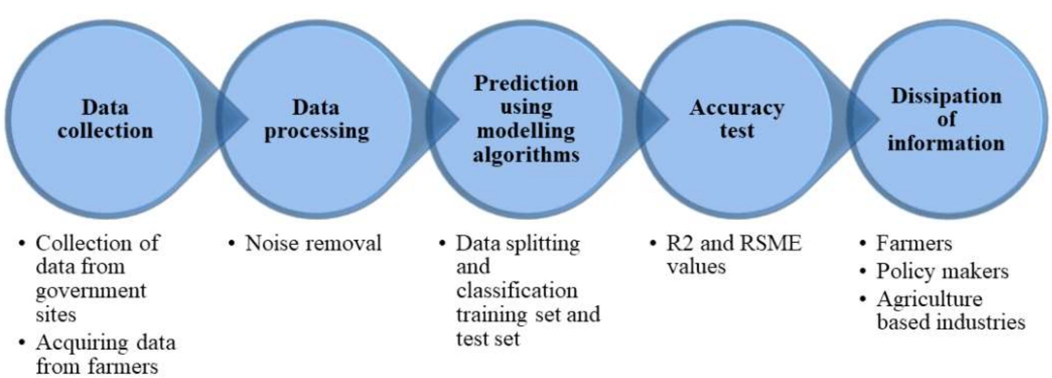

Fig. 1. Flow chart representing crop yield prediction process

Different models based on regression analysis are used in crop yield prediction. Some of these are- shrinkage regression (supports multicollinearity by correcting the level of regression coefficients), ridge regression (maintains all predictors by imposing specific penalties), least absolute shrinkage, and selection operator (LASSO) and Elastic net regression (ENET). The LASSO and ENET models are shrinkage regression models that can manage multicollinearity, regularisation, and variable selection to increase statistical model prediction accuracy and interpretability. For a consistent regression coefficient and automatic variable selection, LASSO eliminators are used. LASSO regression creates models that are simpler and easier to understand since they use a smaller number of factors. Like RIDGE, ENET essentially decreases coefficients and sets some coefficients to zero (like in LASSO). ENET minimises the impact of many elements while not completely eliminating them. Lambda (1) and alpha (0.5) are two parameters used in both LASSO and ENET that must be optimised for the specific study. Wheat yield forecast was done using Lasso (Least absolute shrinkage and selection operator) regression technique (Kumar et al 2019). R2, RMSE, and MAPE parameters were found to be 0.95, 99.27, 2.7 which are statistical parameters for the validation of the accuracy.

Support vector machine: Support vector machine (SVM) is a classification-based regression model in which every information is plotted in a different dimension (Yang et al 2015) and creates a set of hyperplanes with more or infinite coordinates in boundless dimensional space, in which every data variable is mapped to a coordinate. The model learning training is based on linear function but for nonlinear cases, the model employs kernel functions to plot datasets to more dimensional coordinates. SVM is a simple technique where only a few parameters are to be adjusted for the optimization of the model. The model's classifier feature is capable to differentiate between crop and weed plants. Prediction of rice yield in India was performed using SVM-based classification models (Kumar et al 2019). The study employed one against-one multi classification process, training based on kernel function, and cross-validation using k-fold. The dataset used for the modelling included rice yield from 1950 to 2014 year in India and was obtained from the Directorate of Economics and Statistics, Ministry of Agriculture, Government of India. The accuracy of the model for 4-year relative average was obtained when 4-fold cross validation (75.06%) method was used. The accuracy of the model can be further enhanced by revising the training patterns used in the study.

Decision tree: The decision tree is a technique structured to learn a rule common to previous observations. It is a classification-based data mining technique. The model contains nodes (3 kinds), branches (2 kinds), terminal values, strategy, payoff distribution, and the rollback method. The decision node describes a point where a choice is made and is expressed as square. The decision node further extends to a sequence of decision branches and the result of these sequences is represented by each terminal value. The decision tree algorithm includes the following steps: growth of large decision tree followed by reduction (over fitting of data) and finally the tree is pruned (also known as classification tree described) that is used as the base for classification of the datasets. J48 is an open-source algorithm that is a java based implementation and advanced version of C4.5. J48 is designed to build a decision tree on continuous data in the Weka data mining tool. The decision tree technique was used to predict soybean yield in Bhopal district. The productivity was found to be regulated by relative humidity followed by temperature and rainfall (Veenadhar et al 2011).

Random forest: Random forest is a classification data mining algorithm and can build decision trees from both huge and small data to give an efficient prediction (Goel and Abhilasha 2017). The model scans how many generated trees give the same prediction result. The result which gets the maximum score will be shown as output. This algorithm has high versatility and precision in predicting results**.** Random forest algorithm was used to predict cotton yield in Maharashtra using the R package (Prasad et al 2021). Agromet spectral variables and crop yield from 2001 to 2017 were used as input datasets to predict the yield before actual harvest. The model was found to be faster and reliable with a coefficient of determination (R2) value of 69%, 60%, and 39% for September, December, and February months yield. The model was capable of processing large datasets.

Bayesian belief network: Bayesian network is a powerful statistical model which is a clan of algorithmic rules. It deals with uncertain data and the data includes both qualitative and quantitative variables. Bayesian Belief Network is widely used in agricultural datasets.The model is best suited to assess the crop yield taking into consideration the uncertainty of future climate change based on previous climate parameters including temperature, rainfall, and other factors.

Prediction of rice yield in 27 districts of Maharashtra was done using the WEKA tool employing the Bayesian Networks model (Gandhi et al 2016b). In the report, input datasets used for prediction include precipitation, temperature (minimum, average, and maximum), plant's evapotranspiration, area, production, and yield for the Kharif season collected from accessible records of the Government of India.

Comparison of prediction techniques: The present section focuses on some of the earlier studies associated with comparative study of different data mining tools in the area of crop yield prediction in India. The comparison is based onvarious statistical accuracy parameters including coefficient of determination (R2), root-mean-square error (RSME) or F measure (f-score). Different prediction models based on artificial neural network (ANN), stepwise multiple linear regression (SMLR), principal component analysis together with ANN (PCA-ANN), principal component analysis together with SMLR (PCA-SMLR), least absolute shrinkage and selection operator (LASSO), and elastic net (ELNET) were used to predict coconut yield in the west coast of India (Das et al 2019). Weather data (maximum and minimum temperature, relative humidity, wind speed, and solar radiation) and annual coconut yield for the period 2000-2015 were used as input data. Performance of ELNET was found to be best for coconut prediction based on R2 of calibration (2.08) and RSME (2.36). The order of performance of different models was ELNET > LASSO > ANN > SMLR > PCASMLR > PCA-ANN. Among the weather conditions, relative humidity and solar radiation were the major determinants of coconut yield. Another study used five machine learning techniques namely K-Nearest Neighbour (KNN), Naïve Bayes, Multinomial Logistic Regression, Artificial Neural Network (ANN), and Random Forest to predict mustard crop yield in Jammu (Pandith et al 2020). The input data included soil characteristics and yield. Accuracy, recall, precision, specificity, and f-score parameters were analyzed to evaluate the performance of these models. KNN and ANN were found

Table 1. Different techniques used for crop yield prediction in India

Modeling software	Basis of software	Crops	Interaction/Data sets	References
Waikato Environment for Knowledge Analysis (WEKA)	SVM	Rice	Precipitation, minimum temperature, average temperature, maximum temperature	Gandhi et al 2016
	ANN	Rice	Precipitation, Minimum, average and maximum temperature,	Gandhiet al 2016 a, b
	SVM, NaiveBayes	Rice paddy, cotton, sugarcane, groundnut, black gram	Climate variables	Balakrishna et al 2016
	SVM, Naive Bayes, AdaSVM AdaNaive	Crop growing factor	Soil pH, nitrogen, soil depth, temperature	Balar and Patel 2019
	ANN	Cropyield	Minimum and maximum temperature h	Sivanandhini and Prakash 2020
	ANN	Rice	Precipitation, minimum temperature, average temperature	Kulkarni 2018
	LASSO, ENET, PCA, PCA – SMLRANN	Sorghum	Weather (Karnataka)	Sridhara 2020
R package	RF	Cotton	Rainfall, vegetation condition index, standardized precipitation index	Prasad et al 2021
Decision Support System for Agrotechnology Transfer (DSSATT)	SMLR, PCA-SMLR, LASSO,ELNET, ANN, PCA-ANN	Coconut	Maximum and minimum temperature, relative humidity, wind speed and solar radiation	Das et al 2020
SPSS, MS-EXCEL		Potato	Minimum Temperature, Maximum Temperature, Relative humidity, Relative humidity and Wind-Velocity (U.P)	Snehdeep et al 2018
	C 4.5 algorithm	Paddy	Rainfall, Maximum temperature, Minimum temperature, Potential Evapotranspiration	Veenadhari et al 2014
	ANN, SVM, Recursive Partitioning and Regression Tree, Random Forest, Generalized Linear Model	Wheat	Area, Rainfall and Soil type	Gahoi 2019

ANN; Artificial neural network, MLR; Multiple and Stepwise Linear Regression, SLR; Stepwise linear regression, RR; Ridge regression, RF; Random forest, SVM; support vector machine, DSSATT: Decision Support System for Agro technology Transfer to be the best techniques for mustard crop yield predictionwith 0.8405 and 0.8405 f-score, respectively. Naïve Bayes was found to have the lowest accuracy (72.33%) in prediction in the concerned study. Sorghum yield prediction for different districts of Karnataka was done using six multivariate weather-based models viz., least absolute shrinkage and selection operator (LASSO), elastic net (ENET), principal component analysis (PCA) in combination with stepwise multiple linear regression (SMLR), artificial neural network (ANN) alone and in combination with PCA and ridge regression model (Sridhara et al 2020). The datasets included maximum and minimum temperature (Tmax and Tmin °C), relative humidity (RH %), actual evapotranspiration (AET), and rainfall (mm). The study concluded LASSO and ENET were the best models the prediction of sorghum with RSME values varying from 88.99 kg ha^{-1} to 863.87 kg ha^{-1} and 96.89 to 870.98 kg ha^{-1} respectively.

Random forest technique was found to be more accurate to predict wheat prediction in Madhya Pradesh with 94.33% accuracy in comparison to Recursive Partitioning and Regression Tree (93.39%), Generalized Linear Model (91.5%), Neural Network (92.45%) and Support Vector Machine (81.13%) (Gahoi 2019). The studied parameters included area harvested, rainfall, and soil type (Deep Medium Black, Shallow &Medium Black, Alluvial Soil. Mixed Red & Black).

Challenges in the prediction of crop yield: Machine learning techniques are getting more attention in predicting crop yield in the present scenario. But they are associated with some hard challenges (Mayuri and Priya 2018).

Accuracy: The basic hindrance is the requirement of huge accurate previous data as the prediction is based on it. India faces challenges in obtaining and maintaining data from remote agricultural lands. Most of these lands are abandoned during the survey as they fail to meet the minimum hectare of land criteria. To model crop yield, existing methods rely on survey data and other variables linked to crop growth (such as weather and soil conditions). In the United States, where data is copious and of relatively high quality, these approaches have been quite successful. Comprehensive surveys of meteorological factors like the Daymet (Thornton et al 2014) and land cover types like the Cropland Data Layer (Boryan et al 2011) are publically available and considerably aid agricultural production forecast. However, in underdeveloped nations, where accurate yield projections are most needed, weather, soil attributes, and precise land cover data are generally unavailable. On the other hand, remote sensing data is widely available and relatively inexpensive. Existing methods are costly and difficult to expand since they rely on locally gathered survey data. Approaches based on remotely sensed data, such as satellite imaging, have the potential to be a low-cost, high-effective alternative.

Associated factors: The prediction results are not universal. The results would differ based on the seed type, use of fertilizer, and other farmer-driven agricultural practices. The other major challenge is unpredictable arrival of pests and diseases. These may drastically affect the yield due to the absence of enough protection measures in place. So, integration of different factors as input while making prediction is an important aspect.

Limitations in modelling approaches: There are some prediction model-based challenges. The basis of prediction used in the machine learning techniques lacks transparency and interpretability. The recent advancement in the technology of Explainable Artificial Intelligence is helping to overcome the issue. ANN has shortcomings in learning patterns as crop yield is a function of many factors and has inconsistency and unpredictability. Back-propagation neural networks have limitations in selecting large numbers of input variables, hidden layer size, and learning rate. Regression models are widely used in crop prediction and work accurately for linear data but they do not fit fine for complex nonlinear data.

CONCLUSION

Various crop yield modelling techniques are proving promising in agriculture, especially in crop yield prediction. The present study puts together accuracy of different data mining tools used in crop yield prediction and can have future wide applicability. ELNET, LASSO and ANN are some of the best tools for crop yield prediction as concluded from various comparison studies. The prediction is useful for farmers to make a lot of decisions about the best time to sow and harvest seeds, and other decisions such as domestic sustenance and export of crop products. The application of machine learning in yield prediction is expected to grow. An important factor to be considered for using these technologies in agriculture is increasing efficiency and accuracy of the model which can be achieved by integrating ensemble learning algorithms, increasing the input data set and parameter tuning. In the future, we need to develop methodologies focused on models for pre-harvest forecasting, real-time weather parameters and complex agricultural datasets. There is scope for developing tools for soil identification and pest control.

REFERENCES

Balakrishnan N and Muthukumarasamy G 2016. Crop production-ensemble machine learning model for prediction. *International Journal of Computer Science and Software Engineering* **5**(7): 148-153.

Balar M and Patel R 2019. Crop production estimator. *International Journal for Scientific Research & Development* **6**(11): 438-441.

Balkrishna A, Mittal R, Kumar A, Singh D and Arya V 2020. Analysis of policy interventions in agriculture and ICT based mechanistic approach towards sustainability: An Indian perspective. *Asian Journal of Agriculture and Rural Development* **10**(1): 194-213.

Balkrishna A, Sharma G, Sharma N, Kumar P, Mittal R and Parveen R 2021. Global perspective of agriculture systems: From ancient times to the modern era. *Sustainable Agriculture for Food Security: A Global Perspective* 3-46.

Boryan C, Yang Z, Mueller R and Craig M 2011. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. *Geocarto International* **26**(5): 341-358.

Champaneri M, Chachpara D, Chandvidkar C and Rathod M 2020. Crop yield prediction using machine learning. *International Journal of Science and Research* **4**: 1-4.

Dahikar, SS, Rode SV and Deshmukh P 2015. An artificial neural network approach for agricultural crop yield prediction based on various parameters. *International Journal of Advanced Research in Electronics and Communication Engineering* **4**(1): 94-98.

Das B, Nair B, Arunachalam V, Reddy KV, Venkatesh P, Chakraborty D and Desai S 2020. Comparative evaluation of linear and nonlinear weather-based models for coconut yield prediction in the west coast of India. *International Journal of Biometeorology* **64**(7): 1111-1123.

Fente, DN and Singh DK 2018. Weather forecasting using artificial neural network. In *2018 second international conference on inventive communication and computational technologies (ICICCT)* 1757-1761.

Gahoi A 2019. Data analytics for predicting wheat crop production. *Artificial & Computational Intelligence* **1**(1). http://acors.org/ijacoi/VOL1_ISSUE1_06.pdf

Gandhi N, Petkar O, and Armstrong LJ 2016a, July. Rice crop yield prediction using artificial neural networks. In *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development.* 105-110.

Gandhi N, Armstrong LJ and Petkar O 2016b. Proposed decision support system (DSS) for Indian rice crop yield prediction. *Proceedings - 2016 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development* 13-18.

Goel E and Abhilasha E 2017. Random forest: A review. *International Journal of Advanced Research* **7**(1): 251–257.

Gonzalez-Sanchez A, Frausto-Solis J and Ojeda-Bustamante W 2014. Predictive ability of machine learning methods for massive crop yield prediction. *Spanish Journal of Agricultural Research* **12**(2): 313-328.

Jaiswal, B and Agrawal M 2020. Carbon footprints of agriculture sector. In *Carbon Footprints* (pp. 81-99). Springer, Singapore.

Josephine BM, Ramya KR, Rao KVSNR, Kuchibhotla S and Rahamathulla S 2020. Crop yield prediction using machine learning. *International Journal of Scientific and Technology Research* **9**(2): 2102–2106.

Kadbhane SJ and Manekar VL 2020. Development of agro-climatic grape yield model for Nashik Region, Maharashtra, India. *Journal of Agrometeorology* **22**(4): 494-500.

Kulkarni S, Mandal SN, Sharma GS, Mundada MR and Meeradevi 2018. Predictive analysis to improve crop yield using a neural network model. *International Conference on Advances in Computing, Communications and Informatics, ICACCI 2018*, 74–79.

Kumar S, Kumar V and Sharma RK 2019. Rice Yield forecasting using support vector machine. *International Journal of Recent Technology and Engineering* **8**(4): 2588-2593.

Kumar S, Attri SD, and Singh KK 2019. Comparison of lasso and stepwise regression technique for wheat yield prediction. *Journal of Agrometeorology* **21**(2): 188-192.

Kumari V, Agrawal R and Kumar A 2016. Use of ordinal logistic regression in crop yield forecasting. *Mausam* **67**(4): 913-918.

Lata, K,and Chaudhari B 2019. Crop yield pediction using data mining techniques and machine learning models for decision support system. *Journal of Emerging Technologies and Innovative Research (JETIR)* **6**(4): 391-396.

Mahto AK, Alam MA, Biswas R, Ahmad J and Alam SI 2021. Short-term forecasting of agriculture commodities in context of Indian market for sustainable agriculture by using the artificial neural network. *Journal of Food Quality* 1-13. https://doi.org/10.1155/2021/9939906

Mayuri PK and Priya VC 2018. Role of image processing and machine learning techniques in disease recognition, diagnosis and yield prediction of crops: A review. *International Journal of Advanced Research in Computer Science* **9**(2): 788-795.

Pandith V, Kour H, Singh S, Manhas J and Sharma V 2020. Performance evaluation of machine learning techniques for mustard crop yield prediction from soil analysis. *Journal of Scientific Research* **64**(02): 394-398.

Prasad NR, Patel NR and Danodia A 2021. Crop yield prediction in cotton for regional level using random forest approach. *Spatial Information Research* **29**(2): 195-206.

Reddy DJ and Kumar MR 2021. Crop yield prediction using machine learning algorithm. *Proceedings - 5th International Conference on Intelligent Computing and Control Systems, ICICCS* 1466-1470.

Sivanandhini P and Prakash J 2020. Crop yield prediction analysis using feed forward and recurrent neural network. *International Journal of Innovative Science and Research Technology* **5**(5): 1092-1096.

Snehdeep S, Sisodia BVS, Rai VN and Kumar S 2018. Weather forecast models of potato yield using principal componant analysis for Sultanpur district of eastern Uttar Pradesh. *International Journal of Current Microbiology and Applied Sciences* **7**(07): 2000-2006.

Sridhara S, Ramesh N, Gopakkali P, Das B, Venkatappa SD, Sanjivaiah SH and Elansary HO 2020. Weather-based neural network, stepwise linear and sparse regression approach for rabi sorghum yield forecasting of Karnataka, India. *Agronomy* **10**(11): 1-24.

Thornton PE, Thornton MM, Mayer BW, Wilhelmi, N, Wei Y, Devarakonda R and Cook RB 2014. Daymet: Daily surface weather data on a 1-km grid for North America, version 2. Oak Ridge National Lab. (ORNL), Oak Ridge, TN (United States).

Veenadhari S, Misra B and Singh CD 2014. Machine learning approach for forecasting crop yield based on climatic parameters. *International Conference on Computer Communication and Informatics: Ushering in Technologies of Tomorrow, Today, ICCCI* 1-5.

Vishal E, Prasad EH and Gagan V 2021. Prediction and estimation of crop yield using machine learning. *International Journal of Modern Agriculture* **10**(2): 3833-3842.

Vishawaji KP, Sahu PK, Dhekale BS and Mishra P 2016. Modelling and forecasting sugarcane and sugar production in India. *Indian Journal of Economics and Development* **12**(1): 71.

Veenadhari, S, Mishra B and Singh CD 2011. Soybean productivity modelling using decision tree algorithms. *International Journal of Computer Applications* **27**(7): 11-15.

Yang L, He M, Zhang J and Vittal V 2015. Support-vector-machine-enhanced markov model for short-term wind power forecast. *IEEE Transactions on Sustainable Energy* **6**(3): 791-799.

Received 16 January, 2022; Accepted 23 May, 2022