Rice and Potato Yield Prediction Using Artificial Intelligence Techniques

Chiranjit Singha and Kishore C. Swain

Abstract Crop yield prediction during the growing season is important for crop income, insurance projections and even ensuring food security. Yet, modeling crop yield is challenging because of the complexity of the relationships between crop growth and the interrelated predictor variables. This research work employed artificial intelligence (AI) technique for rice and potato crop yield prediction model in the region of Tarakeswar block, Hooghly District, West Bengal, for rice and potato. The major variables used were climatic factors, static soil parameters, available soil nutrient, agricultural practice parameters, farm mechanization, terrain distribution and socioeconomic condition. The analyzed datasets covered 2017 to 2018 seasons and were split into two parts with seventy percent data used for model training and the remaining thirty percent for validation. The mean rice and potato yield obtained from the seventy-farm plot location was about 4.68 t/ha and 18.67 t/ha, whereas the artificial neural networks (ANN) model estimated with 97% accuracy and R^2 value of both the crop is 0.93 and 0.94 with an RMSE of 0.29 t/ha and 1.34 t/ha, respectively. Deep neural networks (DNN) outperformed among the three model, where only support vector machine (SVM) had a sound performance for the training data but low for the validation dataset due to overfitting problem within RMSE and R^2 value. The optimized DNN model produced the highest prediction accuracy 98% for rice and potato crop (RMSE = 0.20 ton/ha and 0.95 t/ha; R^2 = 0.98 and 0.97, respectively), which indicates good correlation between the field-measured crop yield and estimated yield. These adopted methodology for prediction crop yield to provide recommendation to the farmers, decision makers and stakeholders can make farming more efficient and profitable.

Keywords Artificial intelligence · Crop yield · ANN · SVM · DNN

C. Singha · K. C. Swain (✉)
Department of Agricultural Engineering, Institute of Agriculture, Visva-Bharati, Sriniketan, West Bengal 731236, India
e-mail: kishore.swain@visva-bharati.ac.in

P. K. Pattnaik et al. (eds.), *Internet of Things and Analytics for Agriculture, Volume 3*, Studies in Big Data 99, https://doi.org/10.1007/978-981-16-6210-2_9

1 Introduction

Agriculture is backbone of the many developing countries of their socioeconomic development and plays important role in the food management and food security [1]. Climate change, soil variability, water use efficiency, precipitation, humidity, topography, crop practice, weeds, pests and biotic stress, etc., are criteria for monitoring crop yield [2]. Crop yield forecast in precision agriculture study is well thought-out of highly significance for optimization of profit and maximization of crop production. Once the yield is location-specifically projected, the farm inputs such as irrigation, pesticide, mechanization and fertilizers supply could be applied variably according to the accepted soil status and crop requirements. Consequently, it is essential to have tools that facilitate to supervise crop growth and estimation crop yield. Ensuring the management of food demand requires proper monitoring, forecasting and estimating agricultural production for the land parcel [3]. Site-specific crop management (SSCM) dealing with precision agriculture (PA) approach that is measuring, observing and responding to inter- and intra-field spatial variability in soils and crops. Precision agriculture study requires more intensive data collection and information, processing in real time and space to take better crop production decisions of farm inputs, maintaining environmental quality [4]. Whipker and Akridge [5] include growth in demand for both technological advances and information supervision services such as, global positioning system (GPS) auto steering guidance (e.g., Real-Time Kinetic technology), fertilizer, variable rate irrigation and robotics, sprayer controllers and real-time decision making based on sensor networks and remote sensing. A reliable and accurate forecasting model for crop yields is of crucial importance for efficient decision making process in the agricultural sector. Here, widely adopted machine learning algorithms in crop yield prediction such as decision trees RF classification [6], support vector machines (SVMs) [7], naïve Bayes [8], *k* means clustering [9], supervised Kohonen networks (SKNs) [10], eXtreme gradient boosting (XGBoost) [11], light gradient boosting machine (LightGBM) [12], artificial neural networks (ANNs) [13], genetic algorithms (GAs) [14] and ensemble deep neural networks [15] have been used successfully on remotely sensed information in cultivation with high precision. Crop yield modeling is challenging because of the difficulty of the interaction between crop growth and inter- or intra-predictor variables. ANN applied for determining target corn yields using soil properties [16, 17]. According to Noack et al. [18], ANN model is a special network structure with self-adaptive and self-map organizations which contribute to better crop yield estimation as compared to other traditional linear and nonlinear approaches.

The objective of this research work was to deploy three artificial intelligence techniques, SVM, DNN and ANN, as an ordered and monitoring instrument to develop yield prediction model for rice and potato crop in Tarakeswar block, Hooghly, West Bengal, India.

2 Materials and Methods

2.1 Support Vector Machines

Support vector machine (SVM) is a statistical non-parametric, supervised learning approach to classify heterogeneous data that can also be used for regression. SVM is basically designed for binary classification with higher accuracy but can be extended for classification of multiple classes using pair-wise coupling techniques [19]. Main target of the SVM learner has the optimal separation hyperplane (OSH), which is a judgment periphery between classes that reduces classification error in training by having the upper limit margin and afterward generalize to invisible data by kernel functions [20]. According to Vapnik [21], SVM characteristics with nonlinear kernel method is used for model fitting and control the hyperplanes individually grouping sample.

2.2 Artificial Neural Network (ANN)

ANN model applicable between input and output dataset where the data set consists of a linear and highly nonlinear relationship. ANN architecture work out multi-faceted problems with one input layer, one output layer and zero or more hidden layer(s). The ANN model has been used for different crop yield prediction such as, rice [22], wheat [16], potato [23], bitter melon [24], corn and soybean [25] and maize [26]. The popularity of SVM is due to its several promising characteristics, such as the kernel trick and structural risk minimization principle [27]. A robust ANN model relies on the appropriate collection of inputs and of representative training and testing datasets. One hidden layer neural network structure to predict rice and potato yield using input variable on climate, soil and farm practice management [28]. The simulation of the neural network process learning two phases: (I) training the network with known datasets and (II) testing the trained network for model generalization with the validation purpose. The study set of the Levenberg–Marquardt algorithm for chosen to train the selected multilayer perceptron (MLP) whose computations and analyzes inbuilt function into the MATLAB *nn* toolbox. A single layer of output neuron and a single hidden neuron consists of MLPs structure. The MLP function Y essential adjusted across the subsequent linear grouping of multivariate calculation [29] (Eq. 1).

$$Y(x, \omega, \beta) = G_2\left(\sum\nolimits_j \left[\omega_j G_1\left(\sum\nolimits_{ij} \omega_{ij} x_i + \beta_j\right) + \beta\right]\right) \tag{1}$$

where x indicates ith-dimensional involvement trajectory, j represents number of hidden neurons, ω and β are neural weights and biases. The sigmoid tangent activation function G_1 and the linear activation function G_2 [30], are computed as follows (Eq. 2)

$$G_1(\xi) = \frac{2}{1 + \exp(-2\xi)} - 1 \\ G_2(\xi) = \xi \tag{2}$$

where ξ represents weighted sum of evidence from the previous layer of neurons.

The Levenberg–Marquardt backpropagation optimization algorithm [31] coupled with Bayesian regulation used to train MLPs modifies the usual cost function F_e (the sum of squared errors) by considering an additional term, namely the sum of squared neural weights F_ω: (Eq. 3).

$$F = \alpha F_e + \gamma F_\omega \tag{3}$$

where α and γ are objective function parameters automatically set at their optimum values by the Bayesian regularization proposed by MacKay [32]. Bayesian regulation reduces variance errors because the minimization constrains the weights to small values.

Artificial neural network (ANN) structure was preset up to predict rice and potato yield using inputs twenty-one criteria with 10 number of hidden neurons. The target values consider as yield values for the ANN models in each crop. Error backpropagation (EBP) algorithm training by randomizing the network weights and training set order in the ANNs [33]. Under the EBP algorithm, models are instantiated. All the data were normalized to scale of 0–1 for use in the ANN model. The ANN model was trained and tested with measured yield data (obtained from farmers) from 2017 and 2018 (sites were divided into 70% of training versus 30% of testing sites). In a preliminary analysis, the choice of the variables appeared more significant than the number of neurons in the hidden layer. The most adequate combination of variables was thus searched using three hidden neurons, and the optimum number of neurons in the hidden layer was determined afterwards. All variables were tested individually. The one that yielded the highest model performance was selected.

2.3 Deep Neural Network

Generic AI techniques infrequently have difficulties with overfitting that can be resolved through a demanding optimization method in a deep network architecture which overcome the problem of local minima. Backpropagation algorithm improves accuracy through backward and forward optimization. In these courses, suitable activation functions, such as rectified linear units (ReLU) and sigmoid managed the delinquent of vanishing gradients of loss functions, through the backpropagation development. The parameter optimization model used for tuning of our DNN model. The 'activation' function $a()$ is liable for the network scheme's nonlinearity, and designs the real line to nearly subclass of it. We use the linear activation function (ReLU) found to substantially improve performance over earlier alternatives [34], (Eq. 4):

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \tag{4}$$

which is a variant of the ReLU: $a()$ $()x$ x = max 0.

Methods that we used for training neural networks and implemented in the TensorFlow Keras sklearn packages in Jupyter notebook. 70% data is spilt into training and testing data for model validation with 'adam optimization algorithm' MSE loss function [35]. This loss function is reduced through gradient descent backpropagation algorithm with an iterative method [36] (Eq. 5).

$$R = (\gamma - \hat{y})^2 + \lambda\theta^T\theta \tag{5}$$

where $\theta \equiv \text{vec}\left(\beta, \Gamma^1, \Gamma^2, \ldots, \Gamma^L\right)$ and λ denoted tunable hyperparameter. Greater values of λ lead character unbending fits, while values close to zero will typically reason overfitting in higher networks.

2.4 Performance of Model Evaluation

After calibration and validation of the SVM, DNN and ANN model, variables of each farm plot was laid into the model for crop yield prediction per unit area. The performance of the model is evaluated by calculating the root mean square error (RMSE), which gives an estimate of the standard deviation of the residuals (prediction errors), as follows (Eq. 6)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(Z - (Z_i^*)^2\right)}{n}} \tag{6}$$

where Z_i = observed value of the ith observation; Z_i^* is the predicted value of the ith observation; and n = number of points collected. The RMSE tends to place more emphasis on larger errors and, therefore, gives a more conservative measure than the mean absolute error [37].

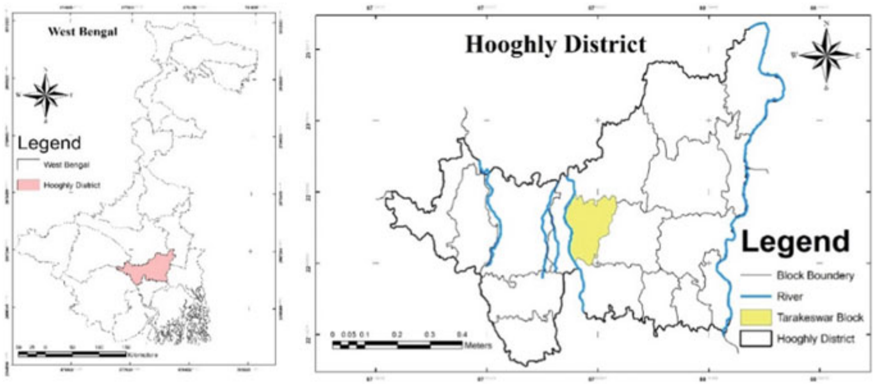

Fig. 1 Study area map

2.5 Study Region

The study region covers with a total area of 119.93 km^2 is located between 22.89° N latitude and 88.02° E longitude and mean elevation of the area is approximately 40 m above sea level (Fig. 1). In the 2011 census, total population of the area was 179,148. The study area is located on old alluvial agro climatic zone. Major growing field crop in *kharif* season is rice and *rabi* season is potato. The cropping intensity of the study area is very high and total land use for crop cultivation 63.85%. The total irrigated area was 11,828 ha [38]. The region characterized by moist subhumid type climate with higher growing period of 180–210 days as result of relatively high rainfall and relative humidity, low PET, warm temperature and low relief. The average annual rainfall of 1350–2500 mm and annual temperature ranging from 10 to 41 °C [39].

2.6 Data Acquisition

In the study area, rice and potato yields are affected by many factors, such as climatic factor, water use efficiency, biotic stress, soil conditions, farm mechanization, terrain distribution and socioeconomic condition. These factors should be measured using appropriate index. A reconnaissance survey of the study area was made in advance of the farming zone for total twenty-one factor such as, soil pH, electrical conductivity (EC), soil organic carbon (SOC), soil texture (ST), available nitrogen (N), available phosphorus (P), available potassium (K), available zinc (Zn), seed rate, mechanization level, irrigation, drainage, pesticide rate, pest affected, source of irrigation, FYM uses, farmer status, NDVI for rice and SAVI or potato, precipitation, temperature, slope and elevation. Various open-source spatial data and secondary information were collected for the proposed crop yield prediction analysis. The elevation and slope degree map of the study area has been produced from the digital elevation model (DEM)—spatial resolution of 30×30 m extracted from SRTMGL1_003 in Google Earth Engine code editor. Climatic data TRMM monthly precipitation and temperature data MOD11A1in degree in acquired by the GEE API.

2.7 Image Processing

In our study, GEE utilized the spatial resolution of 10 m Sentinel-2B Multi-Spectral Instrument, Level-1C, Descending direction and orbit number 33, images with filter metadata for rice 2017-10-01 to 2017-10-30 and for potato 2018-01-15 to 2018-02-15, less than 5% cloud cover images using for rice growing season NDVI and potato season SAVI analysis in the study area [40, 41]; (Eqs. 7 and 8).

$$\mathrm{NDVI} = (\mathrm{NIR}{-}\mathrm{RED})/(\mathrm{NIR} + \mathrm{RED}) \tag{7}$$

$$\mathrm{SAVI} = 1.5\,(\mathrm{NIR} - \mathrm{RED})/(\mathrm{NIR} + \mathrm{RED} + 0.5) \tag{8}$$

where NIR = near infrared band (band 8), R = red band (band 4).

Vegetation indices (Vis) were verified to be one of the maximum effective indicators of crop growing conditions in the study area. NDVI and SAVI have been widely used in crop monitoring and crop yield applications [42]. We trained the classifiers based on a subset of the farmers' declarations. Ground truth data acquired via field inspection were used to develop the models. All parameters are well-defined for reclassifying raster datasets then the extracted value for further analysis using ArcGIS 10.5 software.

2.8 Yield Data

Average crop yield information and non-spatial attributive data of each and every farm plot was derived from field survey during post-harvest periods using GPS ground truth point coordinate order. For yield map generation, kriging interpolation was selected because it is a non-biased method for predicting the values of criteria between the data points assessed using ArcGIS spatial analyst tools with natural breaks jeans method.

2.9 Soil Nutrient Analysis

Between the years of 2017 and 2018, a soil survey was conducted in the study area, resulting in a detailed within each agricultural farm site; seventy soil samples were collected at five evenly distributed points and then mixed thoroughly to obtain a representative sample according to procedures laid out in the Soil Survey Manuals [43]. Soil samples were collected at depths of 0–30 cm from the study area with the special soil auger system. All sampling positions were located by GPS measurements (GPS III Plus, Garmin, Olathe, Kansas, USA). All seventy soil samples were air-dried, crushed, and then passed through a 2.0 mm sieve and the resulting fine earth (<2.0 mm) was retained for further analysis. Measured soil chemical properties included pH (in water, soil/solution ratio of 1:2.5); available nitrogen, determined using Kjehldahl method [44]; available phosphorus (P), determined using the Olsen method [45]; organic carbon (SOC), determined by Walkely and Black method [46]; available zinc, determined by DTPA method with estimation by Perkin Elmer Atomic Absorption Spectrophotometer (AAS) in ppm. Percentages of sand (>50 μm), silt (2–50 μm) and clay (<2 μm) were determined and used to identify the textural class from the textural USDA triangle using hydrometer [47].

3 Results and Discussion

3.1 Accuracy of Yield Prediction

3.1.1 SVM

SVM-based classification models were used for the yield prediction of rice and potato crop. Experiments have been conducted involving one-against-one multi-classification method, *k*-fold cross validation and polynomial kernel function for SVM training with the result 95% accuracy level. The model used the residuals values to check model performance and reviewed after training a model, based on the difference between the predicted and true responses in terms of the trend of regression models. After training a regression model, check the predicted response versus record number. Then SVM cross validate results verified the prediction errors for investigating the predicted and true responses without using the corresponding observation. The SVM model produced the least prediction accuracy for rice and potato (RMSE = 1.09 ton/ha and 5.59 t/ha) crop, respectively (Tables 1, 2 and 3).

Table 1 Area statistics for rice yield measured and predicted

Rice t/ha	Measured		Predicted	
	Area in ha	Area in %	Area in ha	Area in %
<3	17.28	6.15	23.63	8.40
3–4	42.45	15.10	80.43	28.61
4–5	177.01	62.96	97.62	34.72
>5	44.38	15.79	79.45	28.26
Total	281.12	100.00	281.12	100.00

Table 2 Area statistics for potato yield measured and predicted

Potato t/ha	Measured		Predicted	
	Area in ha	Area in %	Area in ha	Area in %
<10	5.38	1.91	51.06	18.16
10–15	63.17	22.47	62.12	22.10
15–20	136.99	48.73	109.06	38.80
>25	75.58	26.88	58.87	20.94
Total	281.12	100.00	281.12	100.00

Table 3 Yield prediction performances of SVM, ANN and DNN model

Model	Accuracy %		RMSE		R^2		Mean predicted yield t/ha	
	Rice	Potato	Rice	Potato	Rice	Potato	Rice	Potato
SVM	95	95	1.09	5.59	0.89	0.90	4.01	18.15
ANN	97	97	0.29	1.342	0.93	0.94	4.74	18.28
DNN	98	98	0.20	0.95	0.98	0.97	4.98	26.8

Note SVM—support vector machine; ANN—artificial neural network; DNN—deep neural network; RMSE—root mean square error

3.1.2 ANN

The supervised ANN was trained with the twenty-one input predictor variables and output yield classes with ten hidden neurons multi-perceptions layer. The best results were structured along with optimum parameters of the artificial neural network modeling for estimating crop yield. In order to test the ability of the neural networks, cross validation was used by leaving 30% of all samples randomly so that after training on the 70% samples, the prediction was verified on this set (Fig. 2, Table 3). In the other hand, ANN model performance of the training, testing and validation performance showing in regression plot gave better agreement than other models (Fig. 2) with 97% accuracy, and R^2 value of both the crop are 0.93 and 0.94, respectively. The best overall results for the prediction of rice and potato yield in cross validation and independent validation were obtained from the ANN

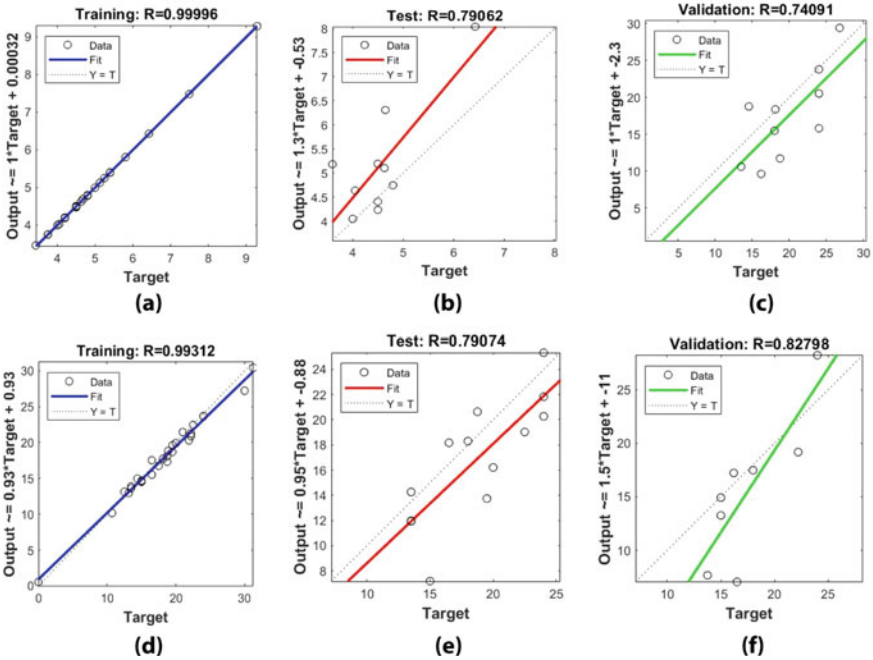

Fig. 2 ANN model performance in training testing and validation sample **a**, **b**, **c** rice; **d**, **e**, **f** potato

networks for the prediction of the low-yield category. The accuracy of prediction reached 97% for both cross validation and independent validation (Table 3). Yield maps the distinction finding between measured and predicted yield is revealed (Fig. 3), where the predicted yield is classified into four groups for rice, as high yield (>5 t/ha), moderate yield (4–5 t/ha), marginal yield (3–4 t/ha) and low yield (<3 t/ha); same way for potato yield classes as high yield (>20 t/ha), moderate yield (15–20 t/ha), marginal yield (10–15 t/ha) and low yield (<10 t/ha) built on four equal class of the yield datasets (Tables 1 and 2). High spatial similarity between the measured and the predicted yield for both the crops (Figs. 3 and 4).

Matsumura et al. [48], reported a close relationship between the predicted yield and the measured yield for maize cultivation in Jilin, China, where the fertilizer and climate variable as a good predictor. Papageorgiou et al. [49] presented that yield classification of field into four different yield categories based on combination of superior predictor variable such as soil, climate and vegetation indices, terrain distribution, local practice and socioeconomic. Uno et al. [50] have utilized remote sensing vegetation indices as crop parameters for predict yield maps similar to our study, where vegetation indices are used, NDVI for rice and SAVI for potato crop as a predictor for crop yield estimation. Area statistics were derived for rice and potato cultivation in *Tarakeswar* block, Hooghly District for different crops yields,

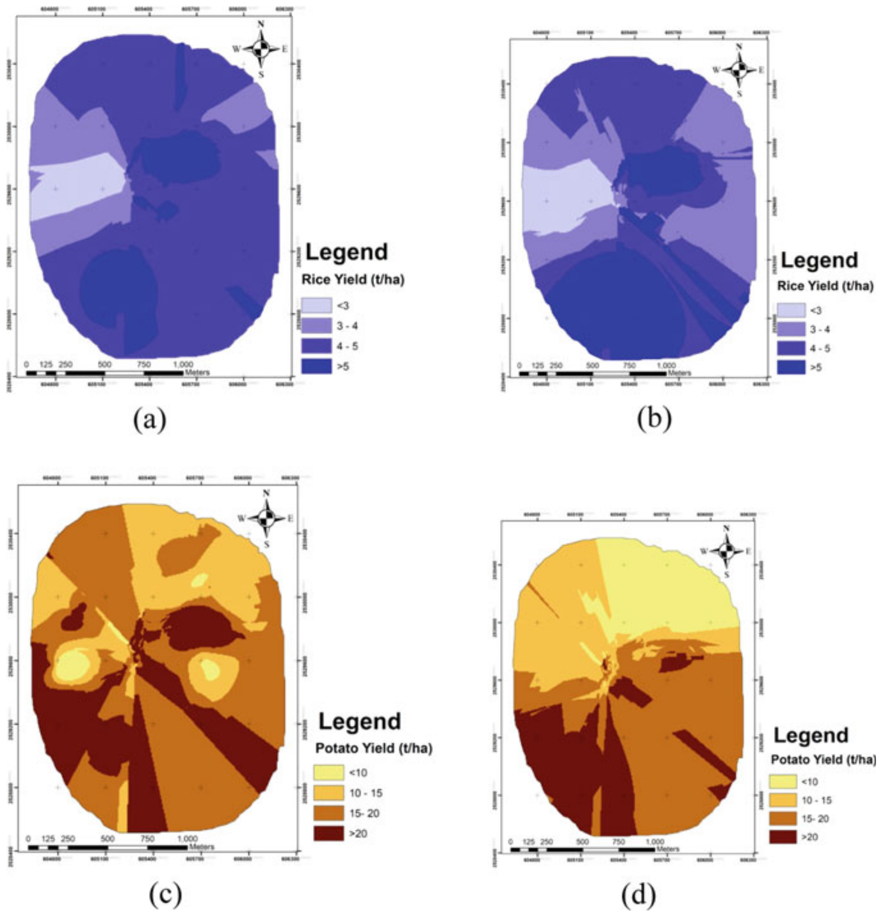

Fig. 3 Crop yield maps estimated by ANN model for rice and potato **a** rice measured **b** rice predicted **c** potato measured **d** potato predicted

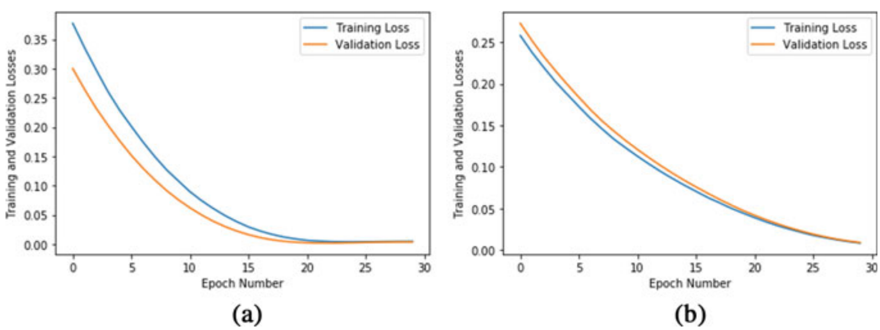

Fig. 4 Model loss progression during training/validation for **a** rice and **b** potato crop

the total area under rice and potato cultivation was also estimated as 281.12 ha (Fig. 3; Tables 1 and 2).

3.1.3 DNN

Srivastava et al. [51] reported that the DNN model output result accuracy improved for crop yield prediction depending upon by fine tuning of hyperparameters optimization level, activation function, layer structure, loss function, intensive optimizer and drop-out ratio. The optimized DNN model produced the highest prediction accuracy for rice and potato (RMSE = 0.20 ton/ha and 0.95 t/ha) with best 98% accuracy for both the crops, respectively (Table 3). Computation power of DNN model is very high due to optimizing the model in local minima within the loss surface. One set of mini batches containing the entire dataset is called 'epoch.' Our research work has 30 epochs and batch size 25 to model progression during training validation process for potato and rice crop yield prediction 26.8 and 4.98 t/ha in the study area with the good predictor for yield (Fig. 4).

4 Discussion

These adopted methodology for prediction of rice and potato yield help to provide recommendation to the farmers, decision makers and stakeholders, to take decisions that can make their farming more efficient and profitable [52]. The mean rice and potato yield obtained from the farm plot location was about 4.68 and 18.67 t/ha, whereas the ANN model estimated average rice and potato was 4.74 and 18.28 t/ha. The R^2 value was found to be 0.933 (93.3%) and 0.941 (94.1%) with an RMSE of 0.29 and 1.34 t/ha and, which indicates that there is good correlation between the field-measured crop yield and estimated yield. Deep neural networks (DNN) outperformed among the three model; where only support vector machine (SVM) had a sound performance for the training data but low for the validation dataset due to overfitting problem within RMSE and R^2 value. DNN model was very well predicting crop yield with low RMSE for the validation dataset nearly for the rice crop (4.98 t/ha) and potato crop (26.98 t/ha) of their respective average yield values. For validation, estimated crop yield values obtained from the model were compared with the field yield value. Training and validation of the model derived the best combination of parameters for estimation of rice and potato crop yield. The scatter plot of the ANN for the predicted versus actual rice and potato yields of training, testing and validation showed better agreement in models' estimation. Deliberating this content by field examination with the local farmer, and it was exposed that although the low yield zones was acidic soil, low organic carbon, high electrical conductivity, low mechanization level and pest affected issue was accountable for the low yield class in the affected zones.

5 Conclusion

A powerful AI technique, integration of SVM, ANN and DNN, provides an effective tool to crop yield prediction and assesses the contribution of each aspect to the target crop yield. Non-influencing factors are adjusted by the weights of the ANN. Two other AI classifiers, SVM and DNN, were shown to be powerful for the classification and crop yield prediction; however, the remarkable results of AI on the agricultural sector enhance in precession agriculture methodologies. Hence, this tool helps the farmers, decision makers and stakeholders take decisions that can make their farming more efficient and profitable. This research tries to establish an intelligent information and crop yield prediction accuracy analysis in sustainable agriculture development. Future studies AI may concentrate on the calibration and trying of this model in macro-level region of organizing data from systematic ground observations, ground sensors, climate, UAV and RADAR remote sensing. Generalized prediction models for diverse crops utilizing parameters development of operational, real-time calculate optimum N rate prediction, crop water stress, like leaf area index, potential evapotranspiration, chlorophyll content, etc., can be developed on same lines for yield forecast.

References

1. Prasad, A.K., Chai, L., Singh, R.P., Kafatos, M.: Crop yield estimation model for IOWA using remote sensing and surface parameters. Elsevier Int. J. Appl. Earth Obs. Geoinf. **8**, 26–33 (2006)
2. Singha, C., Swain, K.C., Swain, S.K.: Best crop rotation selection with GIS-AHP technique using soil nutrient variability. Agriculture **10**, 213 (2020)
3. Bingfang, W., Qiangzi, L.: Crop area estimation using remote sensing on two-stage stratified sampling. Int. Soc. Photogrammetry Remote Sens. (ISPRS) **20**, 12–23 (2004)
4. Mulla, D.J.: Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. Biosys. Eng. **114**(4), 358–371 (2013). https://doi.org/10.1016/j.biosystemseng.2012.08.009
5. Whipker, L.D., Akridge, J.D.: Precision agricultural services dealership survey results. Staff paper. Dept. Agricultural Economics, Purdue University, W. Lafayette, IN, USA (2006)
6. Gyamerah, S.A., Ngare, P., Ikpe, D.: Crop yield probability density forecasting via quantile random forest and Epanechnikov Kernel function (2019). ArXiv: abs/1904.10959
7. Virnodkar, S.S., Pachghare, V.K., Patil, V.C., Jha, S.K.: Remote sensing and machine learning for crop water stress determination in various crops: a critical review. Precision Agric. (2020). https://doi.org/10.1007/s11119-020-09711-9
8. Kaur, S., Kalsi, S.: Analysis of wheat production using Naïve Bayes classifier. Int. J. Comput. Appl. **178**(14), 0975–8887 (2019)
9. Shidnal, S., Latte, M.V., Kapoor, A.: Crop yield prediction: two-tiered machine learning model approach. Int. J. Inf. Technol. (2019). https://doi.org/10.1007/s41870-019-00375-x
10. Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R.L., Mouazen, A.M.: Wheat yield prediction using machine learning and advanced sensing techniques. Comput. Electron. Agric. **121**, 57–65 (2016). https://doi.org/10.1016/j.compag.2015.11.018

11. Zhang, W., Quan, H., Srinivasan, D.: Parallel and reliable probabilistic load forecasting via quantile regression forest and quantile determination. Energy **160**, 810–819 (2018)
12. Cao, J., Zhang, Z., Tao, F., Zhang, L., Luo, Y., Han, J., Li, Z.: Identifying the contributions of multi-source data for winter wheat yield prediction in China. Remote Sens. **12**(5), 750 (2020). https://doi.org/10.3390/rs12050750
13. Kim, N., Ha, K.-J., Park, N.-W., Cho, J., Hong, S., Lee, Y.W.: A comparison between major artificial intelligence models for crop yield prediction: case study of the midwestern United States, 2006–2015. ISPRS Int. J. Geo Inf. **8**(5), 240 (2019). https://doi.org/10.3390/ijgi8050240
14. Martin, C.M.: Crop yield prediction using artificial neural networks and genetic algorithms (2009). http://purl.galileo.usg.edu/uga_etd/martin_charles_m_200912_ms, http://hdl.handle.net/10724/26098
15. Khaki, S., Wang, L.: Crop yield prediction using deep neural networks. Front. Plant Sci. **10**, 621 (2019)
16. Norouzi, M., Ayoubi, S., Jalalian, A., Khademi, H., Dehghani, A.A.: Predicting rainfed wheat quality and quantity by artificial neural network using terrain and soil characteristics. Acta Agriculturae Scandinavica, Sect B—Soil Plant **60**(4), 341–352 (2010)
17. Drummond, S.T., Sudduth, K.A., Joshi, A., Birrell, S.J., Kitchen, N.R.: Statistical and neural methods for site-specific yield prediction. Trans. ASAE **46**(1), 5–14 (2003)
18. Noack, S., Knobloch, A., Etzold, S., Barth, A., Kallmeier, E.: Spatial predictive mapping using artificial neural networks. Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci. **40**(2), 79 (2014)
19. Khobragade, A., Athawale, P., Raguwanshi, M.: Optimization of statistical learning algorithm for crop discrimination using remote sensing data. In: 2015 IEEE International Advance Computing Conference (IACC), pp. 570–574 (2015)
20. Pal, M., Mather, P.M.: Support vector machines for classification in remote sensing. Int. J. Remote Sens. **26**(5), 1007–1011 (2005)
21. Vapnik, V.: Statistical Learning Theory. Wiley, New York, NY, USA (1998)
22. Gandhi, N., Petkar, O., Armstrong, L.J.: Rice crop yield prediction using artificial neural networks. In: IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), Chennai, pp. 105–110 (2016)
23. Fortin, J.G., Anctil, F., Parent, L., et al.: Site-specific early season potato yield forecast by neural network in Eastern Canada. Precision Agric. **12**, 905–923 (2011). https://doi.org/10.1007/s11119-011-9233-6
24. Marizel, B.V., Louella, M., Salenga, M.: Bitter melon crop yield prediction using machine learning algorithm. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **9**(3) (2018). https://doi.org/10.14569/IJACSA.2018.090301
25. Kross, A., Znoj, E., Callegari, D., Kaur, G., Sunohara, M., van Vliet, L., McNairn, H.: Evaluation of an artificial neural network approach for prediction of corn and soybean yield. In: Proceedings of the 14th International Conference on Precision Agriculture, Montreal, QC, Canada, pp. 24–27 (2018)
26. Adisa, O.M., Botai, J.O., Adeola, A.M., Hassen, A., Botai, C.M., Darkey, D., Tesfamariam, E.: Application of artificial neural network for predicting maize production in South Africa. Sustainability **11**, 1145 (2019)
27. Vapnik, V.N.: An overview of statistical learning theory. IEEE Trans. Neural Netw. **10**(5), 988–999 (1999)
28. Liu, J., Goering, C.E., Tian, L.: A neural network for setting target corn yields. Trans. ASAE **44**, 705 (2001). https://doi.org/10.13031/2013.6097
29. Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. Neural Netw. **2**(5), 359–366 (1989)
30. Yonaba, H., Anctil, F., Fortin, V.: Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting. J. Hydrol. Eng. **15**(4), 275–283 (2010)
31. Coulibaly, P., Anctil, F., Bobée, B.: Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. J. Hydrol. **230**, 244–257 (2000)

32. MacKay, D.J.C.: Bayesian interpolation. Neural Comput. **4**, 415–447 (1992)
33. Haykin, S.: Neural Networks: A Comprehensive Foundation, 2nd edn. Prentice Hall, Upper Saddle River, NJ (1999)
34. Nair, V., Hinton, G.E.: Rectified linear units improve restricted Boltzmann machines. In: Proceedings 27th International Conference on Machine Learning (ICML-10), pp. 807–814 (2010)
35. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization (2014). arXiv [preprint] arXiv:1412.6980
36. Crane-Droesch, A.: Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. Environ. Res. Lett. **13**, 114003 (2018)
37. Webster, R., Oliver, M.A.: Geostatistics for Environmental Scientists. John Wiley and Sons, Brisbane, Australia (2001)
38. UNDP.: District human development report (HDR): Hooghly (2011). Development and Planning Department Government of West Bengal (2011)
39. Swain, K.C., Zaman, Q., Jayasuriya, H.P.W., Fang, J.: Estimation of rice yield and protein content using remote sensing images acquired by radio controlled unmanned helicopter. 2008 Providence, Rhode Island, June 29–July 2, 2008 (2008)
40. Huete, A.R.: A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. **25**, 295–309 (1988)
41. Tucker, C.J.: Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. **8**, 127–150 (1979)
42. Sitokonstantinou, V., Papoutsis, I., Kontoes, C., Arnal, A., Andrés, A.P., Zurbano, J.A.: Scalable parcel-based crop identification scheme using sentinel-2 data time-series for the monitoring of the common agricultural policy. Remote Sens. **10**(6), 911 (2018). https://doi.org/10.3390/rs10060911
43. Soil Survey Division Staff.: Soil. In: Survey Manual (Indian Print), Hb. No. IS, USDA, Washington, D.C (2000)
44. Subbaiah, B.V., Asija, G.L.: A rapid procedure for determination of available nitrogen in soil. Curr. Sci. **25**, 259–260 (1965)
45. Olsen, S.R., Cole, C.V., Watanabe, F.S., Dean, L.A.: Estimation of available phosphorus in soils by extraction with sodium bicarbonate. USDA Circular 939. Washington D.C (1954)
46. Walkley, A., Black, L.A.: An examination of datjareff method for determining soil organic matter: and a proposed modification of the chromic acid titration method. Soil Sci. **37**, 29–38 (1934)
47. USDA.: Agricultural research service, soil texture classification. Department of Biological System Engineering, Washington State University, USA (1980)
48. Matsumura, K., Gaitan, C.F., Sugimoto, K., Cannon, A.J.: Maize yield forecasting by linear regression and artificial neural networks in Jilin, China. J. Agric. Sci. **153**, 399–410 (2015)
49. Papageorgiou, E.I., Aggelopoulou, K.D., Gemtos, T.A., Nanos, G.D.: Yield prediction in apples using fuzzy cognitive map learning approach. Comput. Electron. Agri. **91**, 19–29 (2013)
50. Uno, Y., Prasher, S.O., Lacroix, R., Goel, P.K., Karimi, Y., Viau, A., Patel, R.M.: Artificial neural networks to predict corn yield from compact airborne spectrographic imager data. Comput. Electron. Agri. **47**(2), 149–161 (2005)
51. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**, 1929–1958 (2014)
52. Singha, C., Swain, K.C., Nayak, L.: Flood susceptibility mapping through the GIS-AHP technique using the cloud. ISPRS Int. J. of Geo-Inform. **9**(12), 720 (2020)