



Short Communication

Rice yield prediction using hybrid PSACO algorithm based on back-propagation neural network

RAKHEE^{1*}, ARCHANA SINGH¹ and MAMTA MITTAL²

¹Amity University, Noida, India

²DSEU, Dwarka, New Delhi, India

*Corresponding author email : rakheesharma234@gmail.com

Understanding the importance of yield prediction in precision farming alarm the farmers about the future scenario and thereby prepare them to take necessary action to save the crops. There are multiple factors which affect crop growth and yield production such as weather parameters, pest-climate changes, impact of fertilizers, the seeds type etc. The dependencies of the yield are thus a multi-dimensional problem. A growing body of literature has investigated for prediction with high dimensional data (Ezzat et al. 2017, Rakhee et al. 2018, Chen et al. 2020). Reduction to low dimension data and at the same time preserving the important features is a real challenge researcher are facing. Subset of original set of variables which helps in improving the accuracy of the prediction model is obtained through feature selection process (Cadima et al. 2001). In recent years, much work on feature selection is carried out on nature's inspired metaheuristic phenomenon such as Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Gravitational Search Algorithm (GSA) etc. In the literature there are several algorithms focusing on features selection/reduction in agricultural pest/crops (Kashef et al. 2014). Rice yield prediction using different algorithms were evaluated and discussed in recent times as well (Karmokar et al. 2020, Nain et al. 2021). Providing an optimal feature set before a problem of a huge dimension set is an important issue in the feature selection process. Moreover, the quality of selected features has a great impact on prediction problems. Thus, reducing the dimension of features set and at the same time producing the optimal subset for prediction problem is a challenging task. Therefore, the conflict between these two issues generate a need for providing a good search method for near optimal feature subset for the given prediction problem. The present paper investigates how the hybrid of existing algorithms works to improve the efficiency of prediction models by selecting optimal features. The rice crop yield (t/ha) from

Kanpur located in India has been used. In this study yield data from 1971-2019 (procured from Directorate of Economics and Statistics, Department of Agricultural, Cooperation and Farmers Welfare, India) along with weekly (23-32 SMW -kharif season) weather variables (procured from Indian Meteorological Department) viz. maximum temperature (°C), minimum temperature (°C), morning and evening time relative humidity and rainfall (mm) has been used. The data has been subdivided into two sets namely training set (1971-2014) and testing set (2015-2019).

Generation of weather indices

To study the effects of weather variables on yield, indices were generated mathematically and expressed as follows:

$$Z_{ij} = \sum_{t=1}^n r_{it}^j X_{it} \quad \text{and} \quad Z_{ii't} = \sum_{t=1}^n r_{ii't}^j X_{it} X_{i't} \quad (1)$$

Where, r_{it} is correlation coefficient of yield with i^{th} weather variable in t^{th} week and $r_{ii't}$ is correlation coefficient of yield with product of i^{th} and i'^{th} weather variable in t^{th} week. X are weather variables.

Particle swarm optimization (PSO)

In computational world, PSO optimizes the given problem by improving the solutions at each iteration. To seek best solution a particle moves in the direction of particle best (pbest) and global best (gbest) solutions expressed as:

$$pbest(i, t) = \arg \min_{k=1, \dots, t} [f(P_i(k))], i \in \{1, 2, \dots, N_p\} \quad (2)$$

Table 1: Features selected by different algorithms

No.	Algorithm	Features Selected
1	PSO	9 $Z_{20}, Z_{50}, Z_{250}, Z_{350}, Z_{11}, Z_{131}, Z_{231}, Z_{251}, Z_{341}$
2	ACO	5 $Z_{230}, Z_{250}, Z_{350}, Z_{11}, Z_{21}$
3	HPSACO	6 $Z_{20}, Z_{250}, Z_{11}, Z_{21}, Z_{231}, Z_{341}$

Table 2: Model architecture for the prediction of rice yield-Kanpur

No.	Structure of Network	Models
1	9-7-1	PSO-BNN1
2	9-6-1	PSO-BNN2
3	9-5-1	PSO-BNN3
4	5-7-1	ACO-BNN1
5	5-6-1	ACO-BNN2
6	5-5-1	ACO-BNN3
7	6-7-1	HPSACO-BNN1
8	6-6-1	HPSACO-BNN2
9	6-5-1	HPSACO-BNN3

$$gbest(i, t) = \underset{k=1, \dots, t}{\arg \min}_{i=1, \dots, N_p} [f(P_i(k))] \tag{3}$$

where, i denotes the particle index consists of total N_p particles. t is current iteration. f and P are fitness function and position of the particle.

Ant colony optimization (ACO)

Ant colony optimization (ACO) is based on the ability of ant to find the shortest path from their nest to source of food (Uthayakumar *et al.* 2020). Each ant i computes the feasible solution to its current state and move to next state with a probability. The probability of ant to move from one state to another depends on two

Table 3: Comparative performance of the models

Models	NN struc- ture	Training set				Testing set			
		MSE	MAPE	NSE	WI	MSE	MAPE	NSE	WI
PSO-BNN1	9-7-1	0.172	20.186	0.497	0.807	0.502	23.573	0.494	0.790
PSO-BNN2	9-6-1	0.167	20.076	0.549	0.847	0.495	23.447	0.508	0.798
PSO-BNN3	9-5-1	0.183	20.163	0.532	0.836	0.498	23.505	0.495	0.794
ACO-BNN1	5-7-1	0.136	19.127	0.558	0.833	0.488	20.527	0.499	0.759
ACO-BNN2	5-6-1	0.140	19.113	0.563	0.840	0.473	20.477	0.524	0.801
ACO-BNN3	5-5-1	0.131	19.119	0.552	0.837	0.482	20.485	0.502	0.799
HPSACO-BNN1	6-7-1	0.129	19.121	0.604	0.857	0.373	20.187	0.595	0.870
HPSACO-BNN2	6-6-1	0.122	19.109	0.639	0.869	0.366	20.004	0.605	0.823
HPSACO-BNN3	6-5-1	0.123	19.118	0.631	0.861	0.370	20.159	0.600	0.829

factors one is the pheromone level τ_{mn} and another is desirability η_{mn} . The probability of ant i to move from state m to state n is given as follows:

$$p_{mn}(i) = \frac{(\tau_{mn}^\alpha)(\eta_{mn}^\beta)}{\sum_{z \in allowed_x} (\tau_{mz}^\alpha)(\eta_{mz}^\beta)} \tag{4}$$

where τ_{mn} denotes the amount of pheromone deposited by ant i while making transition from m to n and α is the coefficient which control the factor τ_{mn} . η_{mn} stands for desirability to make transition from state m to n and β are the coefficient which control the factor η_{mn} . To reach global optimum solution pheromone needs to update itself at each iteration. The process of updating the pheromone was repeated till optimum solution was found or maximum iteration has reached. Predictions are obtained through back propagation neural network (BNN), The neural network was based on human nervous system. It consists of group of connected units where each connection has

weights associated with it. BNN works with fine tuning of these weights based on error rates obtained in previous iteration each time.

Proposed hybrid algorithm

In this study various weekly weather parameters have been considered for the prediction, weather indices were generated taking care of most relevant interactions of weather parameters for the prediction. The two algorithm viz. ACO and PSO starts simultaneously, at first iteration, position, and velocity of the particle in PSO has been updated. In the same iteration each ant has been assigned to one feature and it visited all features to build a solution. At the end of first iteration the importance of selected subset of each particle (p_{best}) and ant (f_{best}) was evaluated. The best among them would help other to refine their parameters in next iteration according to following rule.

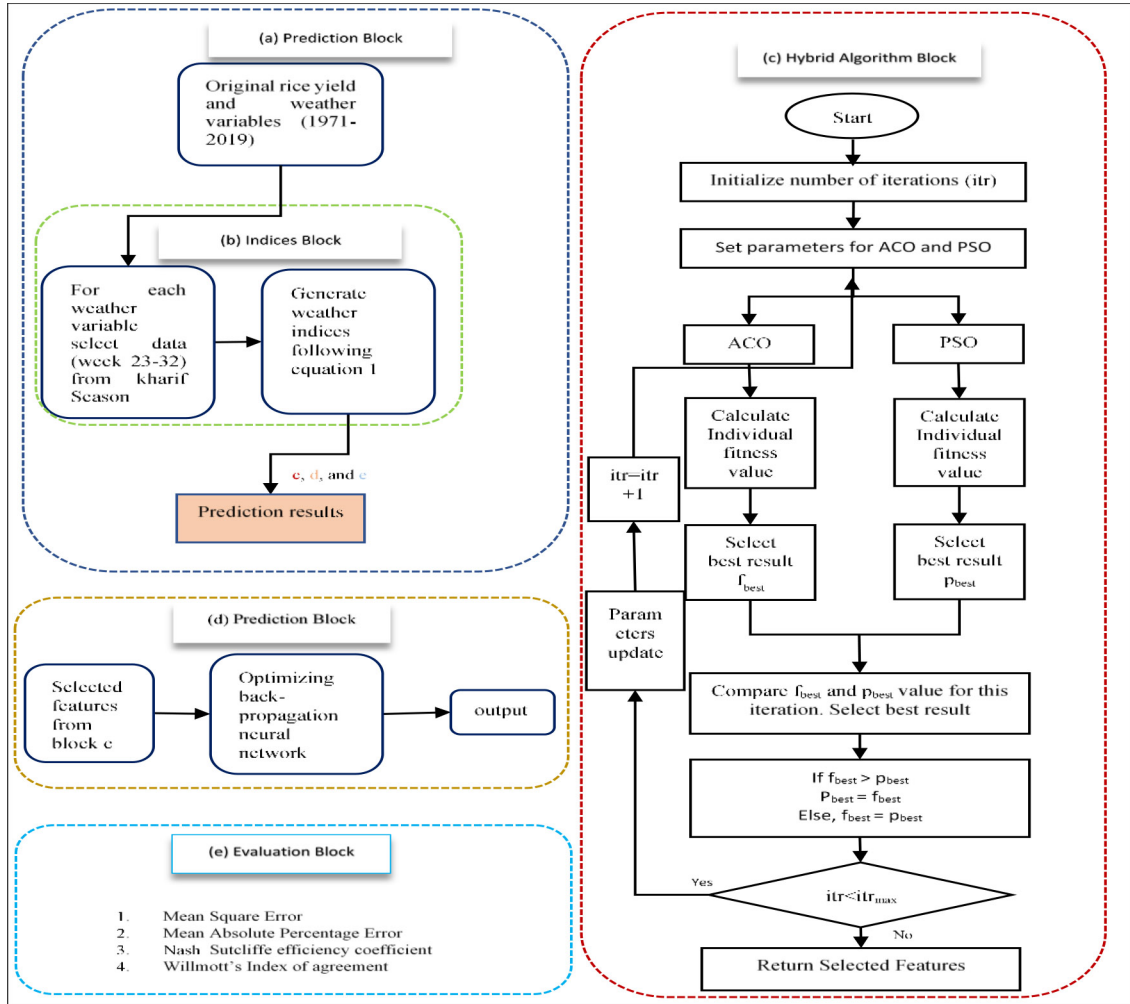


Fig.1: Proposed hybrid algorithm for feature selection and optimization of BNN

$$\text{if } f_{best} > p_{best}$$

$$\text{then, } p_{best} = f_{best}$$

$$\text{else, } f_{best} = p_{best}$$

the best features were selected and considered as independent variables to predict rice yield. Back propagation algorithm is applied on the data for prediction. Levenberg-Marquardt back-propagation neural network was considered as learning algorithm to form a prediction network. For input, hidden and output layer the activation function considered have been linear, hyperbolic tangent and linear respectively, Error function was sum squared error and synaptic function was the dot product. The network has been trained with learning rate of 0.012, 1000 epochs and momentum of 0.36. To reduce the risk of overfitting or trapping in local minima, optimization algorithm was deployed viz. PSO, ACO and HPSACO before training with neural network. The hybrid of algorithms have been shown in Fig 1. These algorithm extracts the features and optimize the network structure to converge to optimal results. Features selected by different optimization algorithm have

been presented in Table 1. The model has been implemented using Matlab 2019a.

From the table, it is evident that through PSO and ACO, 9 and 5 features are extracted respectively, while 6 features are extracted with HPSACO. While making prediction of rice yield with HPSACO algorithm it is evident that total accumulation of minimum temperature, unweighted interaction of minimum temperature and rainfall, weighted accumulation of maximum temperature, weighted accumulation of minimum temperature, weighted interaction of minimum temperature and relative humidity in morning, weighted interaction of relative humidity in morning and evening, would highly impact the rice yield and thus are important in prediction model as inputs. Table 2 illustrate the different structure of network adapted to form prediction models. For each model three architecture of neural network has been reported here. Optimal hidden layer neurons were considered with trial and error method on the prediction model which yielded lowest mean square error. Levenberg-Marquardt algorithm was considered while training BNN as it was proficient than any other algorithm.

Details of the performance of the models have been presented in Table 3. Mean square error (MSE), Mean absolute percentage error (MAPE), Nash-Sutcliffe efficiency coefficient

(NSE) and Willmott's Index of agreement (WI) values for both training and testing sets have been listed. Model PSO-BNN2, ACO-BNN2 and HPSACO-BNN2 can be selected as optimum model for prediction of rice yield with MSE=0.167, 0.140, 0.122 and MAPE= 20.076, 19.113, 19.109 for training data set, and MSE=0.495, 0.473, 0.366 and MAPE= 23.447, 20.477, 20.044 for testing data set. Subsequently it also examines that among three selected architecture HPSACO-BNN2 outperforms the other two models. It is evident from the values of MSE, MAPE, NSE, WI for training and testing set that prediction performance HPSACO-BNN2 was much better than the optimum models PSO-BNN2 and ACO-BNN2. It has the lowest error rate and highest NSE (0.639 for training and 0.605 for testing) and WI (0.869 for training set and 0.823 for testing set) values.

The paper has presented a hybrid approach towards predicting the rice yield of Kanpur located in India. Its hybrid platform includes forming weather indices, feature selection via PSO and ACO and finally predicting the output through BNN. The results obtained from HPSACO-BNN have been compared with PSO-BNN and ACO-BNN via MSE, MAPE, NSE and WI values, respectively. The results have shown that HPSACO-BNN outperforms other algorithm in predicting rice yield. Mean Square error of HPSACO-BNN model with structure (6-6-1) is 0.122 comparing with PSO-BNN (9-6-1) and ACO-BNN (5-6-1) is 0.167 and 0.140, respectively.

ACKNOWLEDGEMENT

Authors are grateful to acknowledge Indian Meteorological Department and Directorate of Economics and Statistics, Ministry of Agriculture, India for providing the data.

Conflict of Interest Statement: The author (s) declares (s) that there is no conflict of interest.

Disclaimer: The contents, opinions and views expressed in the

research article published in Journal of Agrometeorology are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

Publisher's Note: The periodical remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Cadima, J.F.C.L., Jolliffe, I.T. (2001). Variable selection and the interpretation of principal subspaces. *J. Agric. Biol. Environ Stat.*, 6 (1):62-79.
- Chen, Z., Hu, J., Min, G., Zomaya, A.Y., and Ghazawi, T.E. (2020). Towards Accurate Prediction for High-Dimensional and Highly-Variable Cloud Workloads with Deep Learning. *IEEE Trans. Parallel Distrib. Syst.*, 31(4): 923-934.
- Ezzat, A., Wu, M., Li, X., and Kwoh, C. (2017). Drug-target interaction prediction using ensemble learning and dimensionality reduction. *Methods.*, 129:81-88.
- Karmokar, J., Aminulislam, M., Hasssan, M.R. and Billah, M.M. (2020). Impact of Seasonal climatic variability on rice yield in Bangladesh. *J. Agrometeorol.*, 22 (2):165-171.
- Nain, G., Bhardwaj, N., Jaslam, P.K., Dagar, M., and C.S., Anurag. (2021) Rice Yield forecasting using agro-meteorological variables: A Multivariant approach. *J. Agrometeorol.*, 23(1): 100-105.
- Rakhee, Singh, A., Kumar, A. (2018). Weather based Fuzzy regression models for prediction of rice yield. *J. Agrometeorol.*, 20 (4):297-301.
- Uthayakumar, M.N., Shankar, K., and Lakshmanprabu, S.K. (2020). Financial crisis prediction model using ant colony optimization. *Int. J. Inf. Manag.*, 50:538-556.