J. Intell. Syst. 2021; 30:460–469 DE GRUYTER

Research Article

Yinying Cai* and Amit Sharma

Swarm Intelligence Optimization: An Exploration and Application of Machine Learning Technology

https://doi.org/10.1515/jisys-2020-0084
Received Nov 09, 2020; accepted Dec 01, 2020

Abstract: In the agriculture development and growth, the efficient machinery and equipment plays an important role. Various research studies are involved in the implementation of the research and patents to aid the smart agriculture and authors and reviewers that machine leaning technologies are providing the best support for this growth. To explore machine learning technology and machine learning algorithms, the most of the applications are studied based on the swarm intelligence optimization. An optimized V3CFOA-RF model is built through V3CFOA. The algorithm is tested in the data set collected concerning rice pests, later analyzed and compared in detail with other existing algorithms. The research result shows that the model and algorithm proposed are not only more accurate in recognition and prediction, but also solve the time lagging problem to a degree. The model and algorithm helped realize a higher accuracy in crop pest prediction, which ensures a more stable and higher output of rice. Thus they can be employed as an important decision-making instrument in the agricultural production sector.

Keywords: Swarm Intelligence Optimization, Machine Learning Algorithms, V3CFOA, V3CFOA-RF model

1 Introduction

Intelligence Decision Supporting System (IDSS) is widely used today in sectors covering industries, agriculture, transportation, environmental protection, and others [1]. It plays an important part in practices and contributes to the development of humanity. Problems abound in agricultural production: difficulty confronting agricultural pest forecast, crop disease diagnosis, output prediction [2]. After careful study of these problems and machine learning algorithms, a conclusion is reached that suitable swarm intelligence optimization algorithms should be used to better enable machine learning algorithms in solving problems [3]. The current method, with a lack of accuracy or being downright wrong, can elicit serious errors due to innate features of machine learning and swarm intelligence algorithms. The boundaries of machine learning will be factored into two parts: seeking solutions and upgrade in swarm intelligence optimization algorithms so that problems are handily and rationally solvable [4].

Due to the growth in the technology, the more and more researchers have been evolved to uplift the research in the field of efficient agriculture [5]. Various authors have proposed the machine learning techniques to determine the crop sequence and the behaviour of stamps [6]. In addition to this, most of the agriculture relies over the rain, therefore a machine learning model has been proposed by taking environmental data as a training data to decide the perfect future weather conditions for getting proper crop [7] and wool faming of sheep [8]. The crop yield time prediction using several algorithms in the literature has been done by various authors. The some of the algorithms are k-nearest neighbour, artificial neural network and genetic algorithm

***Corresponding Author: Yinying Cai:** Chongqing University of Education, Chongqing 40047, China;
Email: yinyingcai085@gmail.com
Amit Sharma: Jaypee University of Infprmation Technology, Solan, India, 173234; Email: amit.amitsharma90@gmail.com

etc. which have been proposed for the crop yielding time [9]. Also, these algorithms are further extended with the help of pesticides control in the fields. Also, analysis of crop field with soil test is helpful to detect the useful content of soil [10]. Later weather and soil details have been collectively considered together to predict the growth of crop. This mechanism is helpful in case that environmental condition are good but soil is not that much healthy [11].

Research Objectives:
The present paper can be divided into certain objectives which are given as below:

- In the rice crop, the yield of rice crop recognition and prediction time must be known, therefore the proposed algorithm is helpful to obtain the accurate prediction. Along this, the lagging of prediction also solved to some extent with the proposed algorithm.
- To save the rice crop from the insects and other crop infections the pest must be sprayed on the crop which should be timely and in a proper ratio. Therefore, the proposed algorithm predicts the accurate crop pest prediction. Using this prediction, high and stable production of rice has been achieved.
- The proposed algorithm also helpful to design the intelligent agriculture instruments which increased the agriculture production.

Organization:
The rest of the paper is organised in different sections and subsection for better understanding. In section second state of the art of the proposed model is discussed. The proposed model and its flow chart is presented in the third section. Section four is used to represent the experimental setup, extensive results and discussion of the proposed algorithm. The performance comparison is also performed in this section. Section five draws the conclusion and future trends of the proposed work.

2 State of the Art

Machine learning, especially with its ability to learn autonomously, provides new directions for decision-making in agricultural production, and has become an integral instrument in this area [12]. Machine learning classifies data sets collected in real life situation problems into two categories: training set and test set. A decision pattern is generated by studying train set, which is later used in test set to test the efficiency of the pattern [13]. The training and testing carried through real data ensure that problems are solved scientifically and practically [14]. Swarm intelligence optimization algorithms shall be deployed due to the limit of machine learning itself so that better solutions are offered even when facing complex problems [14, 15]. Various researchers have filed patents in this research area. A robot for farming, crop dusting and fertilization has been patented by authors in [16] which is helpful for smart agriculture. In addition to this, crop yielding time, soil test, pest spray and other agricultural functions have been controlled by a robot for agriculture given by the authors in [17].

In literature, various authors have filed patents to aid the agriculture for full pace in the growth of agriculture [18]. Conventional optimization algorithms, including genetic algorithm (GA), artificial bee colony (ABC), particle swarm optimization (PSO), are more complicated when put into actual use because of inevitable deficits of these algorithms, and are always hard to grasp [19, 20]. This makes it imperative to employ a suitable optimization algorithm. Fly fruit optimization (FOA) is a new swarm intelligence optimization algorithm. The creation of the algorithm is inspired by the actual foraging activity of fruit fly [21]. Compared to conventional algorithms, FOA is much easier to use and to grasp in practice thanks to fewer parameters involved [22]. These advantages have made FOA a great tool in solving problems [23]. A study of Chinese and foreign literature in machine learning shows that there are two categories in machine learning: one is to optimize other methods by FOA to improve the performance of algorithms that are still in use; the other is to upgrade FOA itself, including improvements on searching capabilities and expanding boundaries of algorithms [24]. This paper introduces chaotic algorithms to produce a more accurate result on the basis of FOA [25]. The edge of chaotic algorithm over others is that it makes premises based on specific problems, guaranteeing a result is produced in the end [26, 27]. Along with the evolution of computational modelling of swarm intelligence, there has been a consistent expansion in the quantity of research papers revealing the effective use of 3 Swarm Intelligence calculations in various optimization assignments and research issues [28]. Swarm Intelligence standards have been effectively applied in variety of issue areas including function optimization issues, finding optimal paths, planning, structural optimization, and picture and information analysis [29]. The computational modelling of intelligence has been additionally applied to a wide-scope of different areas which includes machine learning, bioinformatics and medical informatics, dynamical frameworks and operational research; they have been even applied in money and business [30].

3 Proposed Model

The proposed algorithm is helpful for the prediction and accuracy in the rice crop yield time prediction, pest spray time. The proposed model is explained in certain steps and explained with the help of flow chart.

3.1 Flow Chart of V3CFOA

To expand search scope and improve efficiency, the original two dimension of FOA is expanded to a third dimension, which increases the rate of convergence [25].

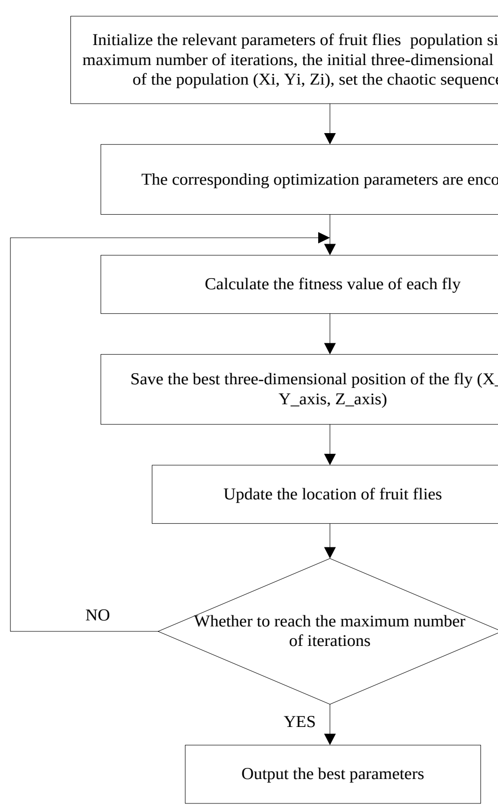

Figure 1: Three dimensional chaotic fruit fly flow chart

Figure 1 is the flow chart of V3CFOA. In the first step, population size has been provided in which the prediction of rice crop takes place. The iteration for training and seed points have been provided for the algorithm as a initialization. The chaotic sequence has been also given in the initialization step. In the next two steps, the optimization parameters have been calculated for the rice prediction and fitness of the output has been calculated for each fly. The coordinate position of the fly has been saved and the location of fly has been updated. This step followed for the complete set of iteration. After completion of iterations the final output of the parameters with best values has been given.

3.2 V3CFOA-RF

The V3CFOA model has been improved by using expansion to third dimension. A third dimension is extended to the previous two dimensions. It creates more room for fruit flies to fly around and optimizes original location for fruit flies. The introduction of chaotic algorithm in FOA adds the odd for it to avoid local optimal [26]. According to the definition of random forest and analysis of algorithms, this paper sets to optimize V3CFOA. First, leverage fruit fly's ability in coordinating its visionary and olfactory senses and its unique capability in seeking the most optimal result. Second, adjust and optimize the amount of decision tree and selecting split attribution which determine the effect of the random forest (RF) model. Third, build an optimized random forest model. Fourth, run a test on data sets concerning crop pests and diseases using k-fold cross validation. V3CFOA model consists the two primal parts: internal parameters optimization and external performance evaluation. In internal parameters optimization, 5-fold cross validation is employed for the dynamic regulation of parameters to make the model more stable and reliable. And in external performance evaluation, the random forest prediction model with the most optimal parameters is used to predict new data sets on crop pests.

4 Results and Discussion

The V3CFOA-RF model is used to test, analyse and compare the data set of insect pests, which will say much about the feasibility and practicability of the proposed algorithm. First, the data set of rice insect pests is selected and the specific description is given so as to provide the experiment with a database. Then the experimental environment parameters are set up according to the characteristics of the algorithm. A compare analysis is made between the algorithm proposed in this paper and conventional algorithms.

4.1 Data on Crop Pests

Statistics used for the paper are mostly from the number of rice stem borer collected in Yushu experimental field in Jilin province from 1989 to 2015. Usually statistics concerning rice pests are influenced by climate, crop growth, weather, etc. A large number of researchers found that the average temperature and average precipitation are the most important factors affecting the number of stem borer.

Therefore data used in this paper are highly representative, including data on the monthly average temperature and average precipitation in May, June, July, August, etc. It is believed that these representative data contribute to the experiment and thus to the predicting of the number of rice stem borer. According to the characteristics of the acquired sample data and test requirements, this paper picks sample data from 1989 to 2010 as training data and data from 2011 to 2015 test data. Part of the training data samples is shown in Table 1.

Figure 2 is a line chart that more obviously reflects the relation between year and the number of stem borer distribution. It suggests that the number of stem borer in most years tend to be stable, and only deviates from the usual stable case once every several years. Through analysis, we can see that the occurrence of insect

Table 1: Part of the training data sample

Attributes	The average temperature in May	Average precipitation in May	The average temperature in June	Average precipitation in June	The average temperature in July	Average precipitation in July	The average temperature in August	Average precipitation in August	Stem borer number
1989	23.90	111.00	27.70	132.00	19.60	41.60	27.10	49.00	22.90
1990	25.30	106.00	27.20	180.00	18.60	93.20	26.90	133.00	22.80
1991	23.70	104.90	26.30	201.00	18.30	44.10	27.00	49.80	24.00
1992	24.50	115.00	23.00	98.00	18.00	49.00	27.50	86.00	22.00
1993	25.20	77.50	28.40	70.30	19.70	76.80	27.80	16.90	23.40
1994	24.10	44.90	26.80	160.00	19.10	45.60	27.80	73.00	24.60
1995	24.50	76.50	28.40	210.00	20.00	14.50	26.50	63.50	22.50
1996	25.20	116.00	26.90	214.00	18.10	107.00	27.90	12.30	22.50
...	...	...	...	...	...	...	...	...	...

pests can be predicted by regression algorithm. V3CFOA is used in order to avoid local optimum caused by local data, and this approach promises a meet of the expectation.

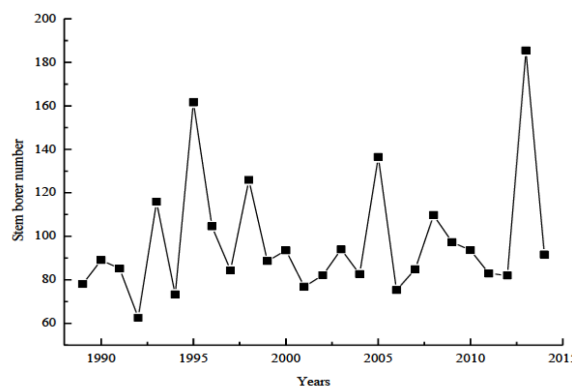

Figure 2: The occurrence amount of stem borers

In order to measure the performance of the proposed method, the following three indexes are compared: RMSE (root mean square error), MAE (mean absolute error) and correlation coefficient R. The three indicators are defined as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(|\hat{y}_i - y_i|\right)^2} \tag{2}$$

$$R = \frac{cov(\hat{y},\ y)}{\sigma_{\hat{y}}\sigma_y} \tag{3}$$

In equation (1) and (2), y_i is actual output, $\hat{y}_i$ is the output generated through the algorithm and N is sample number in training set.

4.2 Experiment Environment and Algorithm Parameter Set

Comparison between and analysis about the algorithm proposed in the paper and other conventional ones are made to show the performance of the model proposed, which is the yardstick that says whether it meets the demand of the experiment or not. According to the analysis of common prediction algorithm models, we selected PSO-RF, FOA-RF and support vector machine (SVM) as contrast objects, and compared the results obtained from each algorithm with that of V3CFOA-RF algorithm, so as to measure the effectiveness of V3CFOA-RF algorithm.

This research adopts the MatlabR2014 development platform, SVM uses the function provided by the LIBSVM toolbox, and the FOA and PSO algorithm models are built. In this paper, the iterative termination condition of FOA and PSO algorithm is the maximum number of iterations. After testing many related parameters, we choose the most suitable value as the parameter of subsequent experiments.

4.3 Experiment Comparison and Data Analysis

In this experiment, we compared the following methods: V3CFOA-RF, FOA-RF, SVM and BP in detail. Through the analysis of a large number of related researches, it is found that grid search is the preferred method in selecting SVM parameters. Therefore, in this paper, SVM parameter selection method based on grid search is applied to realize SVM parameter selection. The training and prediction results using SVM are shown in Figure 3 and Figure 4. They suggest that the fitting effect of SVM in the training set and test set is good. Values from the training set and test set are: 9.6535 and 31.0209 in RMSE, 7.2876 and 25.3158 in MAE, 0.9082 and 0.9083 in correlation coefficients respectively.

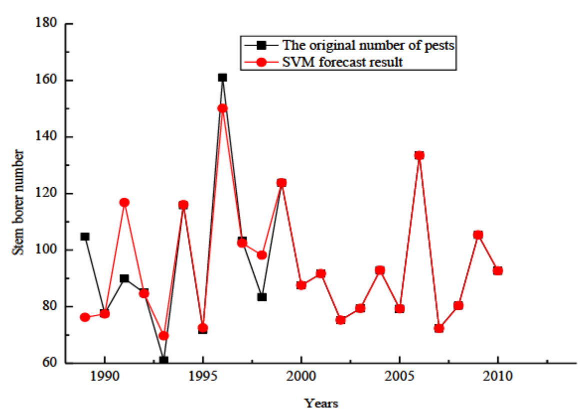

Figure 3: The prediction result of the stem borer number by SVM on training set

To get better prediction results, we use V3CFOA algorithm to search two parameters in RF, get the best parameter value, and then predict the number of pests. Figure 5 and Figure 6 show the fitting results of V3CFOA-RF on the training set and test set, and the prediction results from these two figures show that the method has a high prediction accuracy.

In order to observe the optimization process of V3CFOA parameters and check the optimization ability of V3CFOA, the evolution process of V3CFOA is compared with that of FOA-RF in the paper. Figure 7 is the iterative evolution graph based on V3CFOA method and FOA-RF method. The abscissa represents the iteration number of the algorithm, and the ordinates represent the RMSE results obtained by these two methods.

As can be seen from Figure 7 that in the early stages of iteration, V3CFOA has begun to converge, while FOA-RF has a higher RMSE value, it is until approaching the 200 iteration that FOA-RF starts to converge, while the V3CFOA converges close to the optimal solution and FOA-RF is yet to get the best result even with the maximum number of iterations.

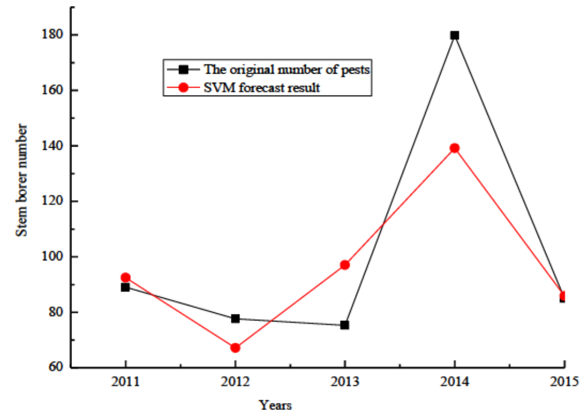

Figure 4: The prediction result of the stem borer number by SVM on the test set

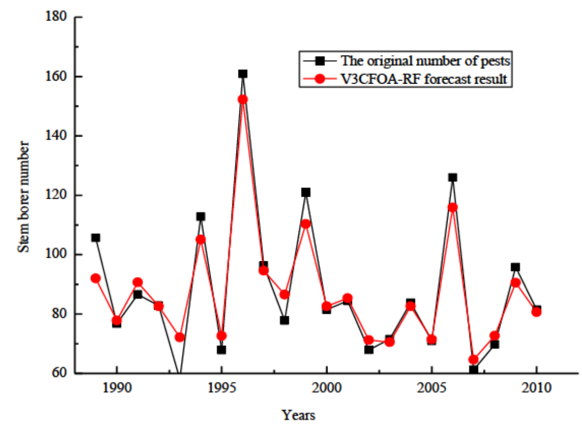

Figure 5: The prediction results of V3CFOA-RF based on the training set

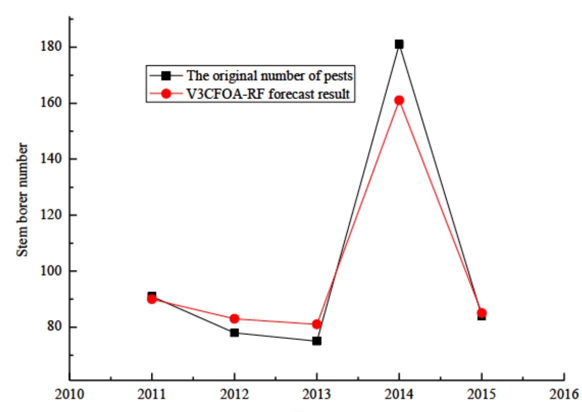

Figure 6: The prediction results of V3CFOA-RF borer based on the test set

Through the analysis of the experimental results, we can see that V3CFOA not only converges faster than FOA, but also has higher reliability and stability than FOA, which indicates that V3CFOA has better performance in RMSE.

The detailed comparison results of the four methods are shown in Table 2. As it shows, V3CFOA-RF is better than PSO-RF, FOA-RF, and SVM measured by all three indicators: RMSE, MAE, and correlation coefficient R. In comparison of the results generated through all these methods, the performance of FOA-RF is slightly inferior to that of V3CFOA-RF, its training set scores 9.3009 in RMSE, 6.3644 in MAE and 0.9639 in correlation

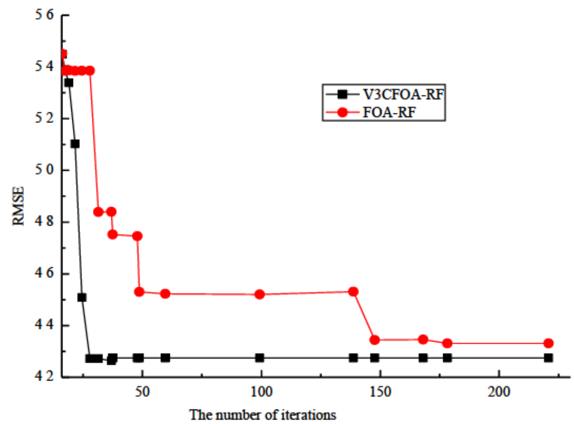

Figure 7: Iterative evolution curve of V3CFOA-RF and FOA-RF

coefficient R respectively, and on the test set, it gets 23.1811 in RMSE, 20.1661 in MAE and 0.9286 in correlation coefficient R. Compared with the random forest method, the SVM result is most atrocious. It is also suggestive that random forest has a stronger generalization ability than the SVM in pest prediction.

Table 2: Comparison results of the four methods

Methods	Training set			Test set		
	RMSE	MAE	R	RMSE	MAE	R
V3CFOA-RF	9.0651	5.6568	0.9770	21.5442	18.4600	0.9513
FOA-RF	9.3009	6.3644	0.9639	23.1811	20.1661	0.9286
PSO-RF	9.5023	7.0645	0.9493	28.8131	23.6671	0.9167
SVM	9.6535	7.2876	0.9082	31.0209	25.3158	0.9083

Meanwhile, during the experiment, we also recorded the CPU usage time of the two methods. It was found that the V3CFOA method has an obvious advantage in time. The whole model training and testing process took only 175 seconds, while the PSO method used 389 seconds. This indicates that, compared with other prediction methods, V3CFOA-RF is more accurate and efficient in prediction, so the algorithm proposed in this paper is more suitable for rice pest prediction.

5 Conclusion and Future Development

This paper is dedicated to the optimization of current machine learning technologies, by studying and utilizing swarm intelligence optimization, to work out intelligent decision-making in agricultural sector grounded on machine learning technologies, and putting into good use of these technologies in actual production situation. V3CFOA-RF, a random forest model of chaotic algorithm developed after FOA algorithm, was put forward to address the problem that parameters affect random forest prediction procedures. Briefly, there are three major parts involved in the study. As a starter, FOA is extended to a third dimension from its previous two dimension. Meanwhile, chaos theory is introduced to initialize the population to avoid local optimum. The result in this stage is an optimized V3CFOA. Following this, V3CFOA is employed to train random forest, and the most optimal algorithm will be obtained. The final step is to put into test the algorithm on rice pest data sets and compare its result with that of other algorithms. The experiment shows that this model has proven to be more accurate in prediction and can more efficiently forecast rice pests, making it a useful tool in agricultural production sector. In future, patents can be filed with respective design perspective to solve the detailed issues in smart agriculture. The advanced learning and training model will further increase the accuracy in the prediction of the rice crop and others.

Funding: None.

References

[1] H. R. Nemati, D. M. Steiger, L. S. Iyer, and R. T. Herschel, “Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing,” *Decision Support Systems*, vol. 33, pp. 143-161, 2002.

[2] B. S. McIntosh, J. C. Ascough II, M. Twery, J. Chew, A. Elmahdi, D. Haase, *et al.*, “Environmental decision support systems (EDSS) development–challenges and best practices,” *Environmental Modelling & Software*, vol. 26, pp. 1389-1402, 2011.

[3] Z. Doshi, S. Nadkarni, R. Agrawal, and N. Shah, “AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms,” in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2019, pp. 1-6.

[4] A. K. Z. Al-Saedi, R. Ghazali, and M. M. Deris, “Improved Multi Join Query Optimization for RDBMS Based on Swarm Intelligence Approaches,” *Advanced Science Letters*, vol. 22, pp. 2847-2851, 2016.

[5] L. Hazard, P. Steyaert, G. Martin, N. Couix, M.-L. Navas, M. Duru, *et al.*, “Mutual learning between researchers and farmers during implementation of scientific principles for sustainable development: the case of biodiversity-based agriculture,” *Sustainability Science*, vol. 13, pp. 517-530, 2018.

[6] R. Johnson and T. Zhang, “Learning nonlinear functions using regularized greedy forest,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, pp. 942-954, 2014.

[7] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70-90, 2018.

[8] S. Shahinfar and L. Kahn, “Machine learning approaches for early prediction of adult wool growth and quality in Australian Merino sheep,” *Computers and electronics in agriculture*, vol. 148, pp. 72-81, 2018.

[9] K. Teeda, N. Vallabhaneni, and T. Sridevi, “Comparative Analysis of Data Mining Models for Crop Yield by Using Rainfall and Soil Attributes,” in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2018, pp. 1176-1180.

[10] F. Borghesan, M. Chioua, and N. F. Thornhill, “Forecast of persistent disturbances using k-nearest neighbour methods,” in *Computer Aided Chemical Engineering*. vol. 44, ed: Elsevier, 2018, pp. 631-636.

[11] A. Shah, A. Dubey, V. Hemnani, D. Gala, and D. Kalbande, “Smart Farming System: Crop Yield Prediction Using Regression Techniques,” in *Proceedings of International Conference on Wireless Communication*, 2018, pp. 49-56.

[12] Y. Abdi, S. Parsa, and Y. Seyfari, “A hybrid one-class rule learning approach based on swarm intelligence for software fault prediction,” *Innovations in Systems and Software Engineering*, vol. 11, pp. 289-301, 2015.

[13] X. He, H. Guan, and J. Qin, “A hybrid wavelet neural network model with mutual information and particle swarm optimization for forecasting monthly rainfall,” *Journal of Hydrology*, vol. 527, pp. 88-100, 2015.

[14] G. Kiruthiga, S. Krishnapriya, V. Karpagambigai, N. Pazhaniraja, and P. VicterPaul, “Survey on swarm intelligence based optimization algorithm,” *International Journal of Applied Engineering Research (IJAER)*, vol. 10, pp. 18445-18457, 2015.

[15] J. Albinati, S. E. Oliveira, F. E. Otero, and G. L. Pappa, “An ant colony-based semi-supervised approach for learning classification rules,” *Swarm Intelligence*, vol. 9, pp. 315-341, 2015.

[16] W. E. Allen and J. P. Sewell, “Method and apparatus for establishing a perimeter defining an area to be traversed by a mobile machine,” ed: Google Patents, 2000.

[17] K. Cavender-Bares and C. C. Bares, “Robotic platform and method for performing multiple functions in agricultural systems,” ed: Google Patents, 2016.

[18] B. P. Hanley, “Flexible agricultural automation,” ed: Google Patents, 2003.

[19] Y. M. B. Ali, “Unsupervised clustering based an adaptive particle swarm optimization algorithm,” *Neural Processing Letters*, vol. 44, pp. 221-244, 2016.

[20] C. Zhang, Y. Yang, Z. Du, and C. Ma, “Particle swarm optimization algorithm based on ontology model to support cloud computing applications,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 7, pp. 633-638, 2016.

[21] P. Guo, W. Cheng, and Y. Wang, “Parallel machine scheduling with step-deteriorating jobs and setup times by a hybrid discrete cuckoo search algorithm,” *Engineering Optimization*, vol. 47, pp. 1564-1585, 2015.

[22] H. Liu, G. Xu, G. Ding, and D. Li, “Integrating opposition-based learning into the evolution equation of bare-bones particle swarm optimization,” *Soft Computing*, vol. 19, pp. 2813-2836, 2015.

[23] R. Forsati, A. Keikha, and M. Shamsfard, “An improved bee colony optimization algorithm with an application to document clustering,” *Neurocomputing*, vol. 159, pp. 9-26, 2015.

[24] S. K. Agarwal, S. Shah, and R. Kumar, "Group based Swarm evolution algorithm (GSEA) driven mental task classifier," *Memetic Computing*, vol. 7, pp. 19-27, 2015.
[25] K. Wang and F. Ye, "Improved Encryption Algorithm of Images Based on Three-Dimensional Chaos," *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, vol. 9, pp. 73-83, 2015.
[26] A. A. Heidari, R. A. Abbaspour, and A. R. Jordehi, "An efficient chaotic water cycle algorithm for optimization tasks," *Neural Computing and Applications*, vol. 28, pp. 57-85, 2017.
[27] X. Yuan, T. Zhang, X. Dai, and L. Wu, "Master–slave model-based parallel chaos optimization algorithm for parameter identification problems," *Nonlinear Dynamics*, vol. 83, pp. 1727-1741, 2016.
[28] Sharma, A., Kumar, R., Talib, M. W. A., Srivastava, S., & Iqbal, R. (2019). Network modelling and computation of quickest path for service-level agreements using biobjective optimization. *International Journal of Distributed Sensor Networks*, 15(10), 1550147719881116.
[29] Poongodi, M., Sharma, A., Vijayakumar, V., Bhardwaj, V., Sharma, A. P., Iqbal, R., & Kumar, R. (2020). Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system. *Computers & Electrical Engineering*, 81, 106527.
[30] Mavrovouniotis, M., Li, C., & Yang, S. (2017). A survey of swarm intelligence for dynamic optimization: Algorithms and applications. *Swarm and Evolutionary Computation*, 33, 1-17.