TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 18, No. 6, December 2020, pp. 3011~3018
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v18i6.15870

Predictions on wheat crop yielding through fuzzy set theory and optimization techniques

Julio Barón Velandia, Norbey Danilo Muñoz Cañón, Brayan Leonardo Sierra Forero
Facultad de ingeniería, Universidad Distrital Francisco José de Caldas, Colombia

Article Info

Article history:

Received Feb 18, 2020
Revised May 25, 2020
Accepted Jun 25, 2020

Keywords:

Agriculture
Crops
Fuzzy set theory
Optimization techniques
Performance
Prediction
Wheat

ABSTRACT

Agricultural field's production is commonly measured through the performance of the crops in terms of sow amount, climatology, and the type of crop, among other. Therefore, prediction on the performance of the crops can aid cultivators to make better informed decisions and help the agricultural field. This research work presents a prediction on wheat crop using the fuzzy set theory and the use of optimization techniques, in both; traditional methods and evolutionary meta-heuristics. The performance prediction in this research has its core on the following parameters: biomass, solar radiation, rainfall, and infield's water extractions. Besides, the needed standards and the efficiency index (EFI) used come from already developed models; such standards include: the root-mean-square error (RMSE), the standard deviation, and the precision percentage. The application of a genetic algorithm on a Takagi-Sugeno system requires and highly precise prediction on wheat cropping; being, 0.005216 the error estimation, and 99.928 the performance percentage.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Brayan Leonardo Sierra Forero,
Facultad de ingeniería,
Universidad Distrital Francisco José de Caldas,
Carrera 7 No. 40B – 53, Bogotá D. C. República de Colombia.
Email: blsierraf@correo.udistrital.edu.co

1. INTRODUCTION

To predict any natural event requires a logical analysis on the occurrence frequency and the nature of the event itself [1]. Therefore, to apply any process aimed to develop a prediction model, researchers must take advantage of those methods that handle uncertainty better than the traditional ones do [2]; this, allow them to create models that would improve the outcomes (depending on the context and goals establishment) and the model's significance and accuracy. Some of those methods suitable for prediction are the fuzzy logic systems, the artificial neural networks, and the adaptive neural-fuzzy systems. Such methods used widely in the prediction of agricultural performance and the estimation of variables and parameters [3] as crop, soils, climatology, the convenience [4-6] or suitability of a specific crop in a determined geographical zone [7, 8], and even, the optimization of processes related to the nature of the systems itself [9, 10]. Furthermore, other methods, specialized on optimization, also allow researchers to create prediction models. Such, are divided into traditional methods (focused on local search and general solution) and heuristic methods (stochastic or probabilistic, and focused on solution or target population) [11]; optimization techniques are algorithms intended to find optimal solutions to specific problems [12]. The latter usually include gradient based algorithms, free gradient algorithms, evolutionary algorithms, and nature-inspired meta-heuristics [13].

A gradient-based algorithm is an iterative method that cycles the target's function gradient information throughout the iterations of the process [13]; the generic Quasi-Newton method has a super-linear convergence rate that separates it from other traditional methods as the gradient descent method (linear convergence rate) and the Newton-Raphson method (squared convergence rate) [14]. Otherwise, genetic algorithms (GA) imply the coding of the target function as bits matrices or character chains that depict chromosomes; the manipulation of those chains by genetic operators; and the selection of suitable solutions to a give problem [13]. In contrast to the other stochastic methods, GAs conceptual development is supported on mathematical theory, greatly on the Scheme Theorem, which conjugates fitness, crossbreeding, and mutation to establish how survival and solution propagation would be affected [11].

Crops' outcome is measured regarding its performance and considering the factors (climatic, biotic, and edaphic) that could affect them; such factors and its constituents impact on the crops is never isolated, on the contrary, it is interdependent [15]. According to Syngenta, sub-factors as solar radiation (growing); rainfall volume and soil (limiting); and plagues and illnesses (reduction), directly affect the crops' performance. Furthermore, biomass is synthesized via biotic organic components in which water intercede as the vehicle for chemical reactions and solar radiation supports the energetic needs of the crop [16]. Therefore, biomass, rainfall volume, solar radiation, and in-field water extractions, will be the sub-factors established to develop the prediction model for wheat crop in this research.

This research presents an evaluation of the predictive model performance through comparison of two configurations based on the fuzzy set theory for the forecasting of a wheat crop yielding: generic Quasi-Newton gradient algorithm (traditional optimization) and GA (heuristic method). This document will define the research method; its configurations and the optimization techniques used; and will present a brief description regarding the APSim data-set used and the data extracted from [17].

2. RELATED RESEARCH

Modified GAs can solve multi target issues where stochastic optimization is a must and can help cultivators to make better decisions regarding the cost/utility ratio on agricultural operations [18]; one of its uses can be seen in the estimation on the change times for medium size macadamia nut crops. As part of the modern agricultural practices, greenhouse cropping requires accurate models for plant growing (and other parameters) under a series of climatic factors. To solve such problem research was carried out through a double GA in which the primary algorithm parameterize the model, while the secondary one defines the initial algorithmic parameters [19].

An optimization method, based on an ACO and on an advanced Process-focused cropping model, and tested on a corn crop in Colorado (USA), was carried out to reduce the water and fertilizer use to its minimum taking into account the bio system's reference [20]. GAs implementation on rice crops would derive on models to improve both, productivity and quality, while optimizing its yielding. On the model proposed for rise cultivation 16 parameters are minded to alter the performance of the cereal drops in Thailand, on this example risk management is included as an outcome and throughout iterations on the risk level over the rice crops the lesser risk solution was attained [21].

Stochastic algorithms can be combined to generate hybrid meta-heuristics. In this case, the particle swarm optimization (PSO), imperialist competitive algorithm (ICA), and support vector regression (SVR) were combined to predict the performance of apricot crops in Abarkuh Yazd (Iran): 61 variables were considered, 18 of those were more influential on the crops' performance according to the use of the hybrid algorithm [22]. Another manner to accurately estimate the performance of a crop is using data and indexes from crop-growing simulation tools. For instance, this research assumed biomass and dosel covers coming from the vegetation indexes on the aqua crop simulation model (FAO), through a PSO, to obtain more accurate estimations for those factors on corn crops. The outcome was validated through an root-mean-square error (RMSE) and compared to other vegetation indexes [23].

3. RESEARCH METHOD

The application of an experimental method allows the adaptation of the environment to obtain an expected result, looking for the characteristics and properties relevant in the experiment [24]. It focuses on the analysis of the set of records that describe crop behavior in terms of critical variables and yield. This section describes the flow of activities carried out to obtain the optimal value for performance. Each subsection represents a configuration as follows: the first corresponds to the configuration of the fuzzy systems and the second to the implementation of optimization techniques. This research used 680 daily compiled datasets regarding wheat crops from the APSim framework:

3.1. Fuzzy set configuration

In MATLAB, two configurations were set; both with the above mentioned four input factors and a single outcome (crop performance) responding to the interaction of the input data. The first configuration is based on Mamdani model with a simple structure of operators "min -max"; 16 fuzzy inference rules are adjusted; and each input and output are assigned with 3 belonging trimf functions: low, medium and high. The second configuration uses the Takagi-Sugeno model, this has extra options regarding implication, adding, and un-fuzzing processes; 10 fuzzy inference rules are adjusted; and each input and output is assigned with 3 belonging gaussmf functions. Figure 1 shows the configuration's definition.

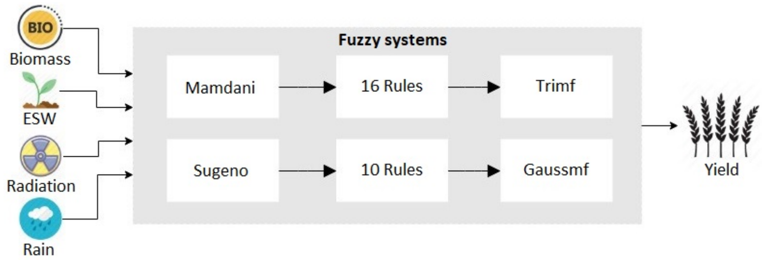

Figure 1. Definition of the two configurations of fuzzy systems

3.2. Optimization techniques

Two optimization algorithms were applied to each configuration defined: the first one is a gradient-based optimization established to find the functions' minimum for each parameter; the fuzzy set is adjusted for each iteration based on the mean-square error (MSE). The second one is GAs optimization intended to create the ranges for each belonging function according to the parameter dataset; those ranges are adjusted according to the difference operation between real and predicted data. This algorithm, applied on the fuzzy system, requires the maximum setting on the values iterations and optimization time. Figure 2 shows how these optimization algorithms work.

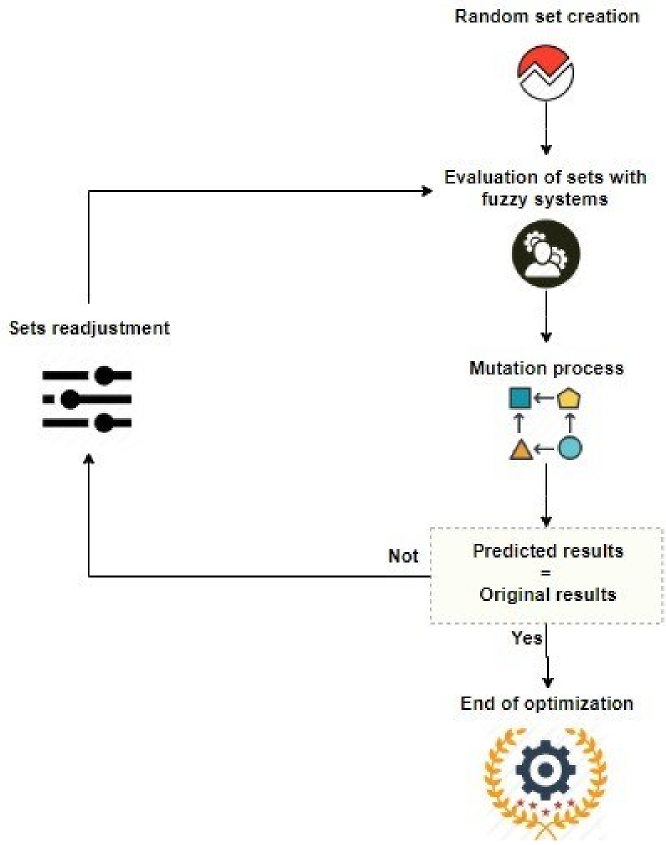

Figure 2. General operation of optimization algorithms

Figure 3 describes the process flow that represents the experimental method applied to configure fuzzy systems and optimization algorithms. The procedure begins with the collection and preprocessing of the data; next, membership functions are designed and inference rules are configured; afterwards the fuzzy systems and the respective optimization are implemented using the algorithms defined to finally obtain the optimal value of the performance through the validation and comparison of the results.

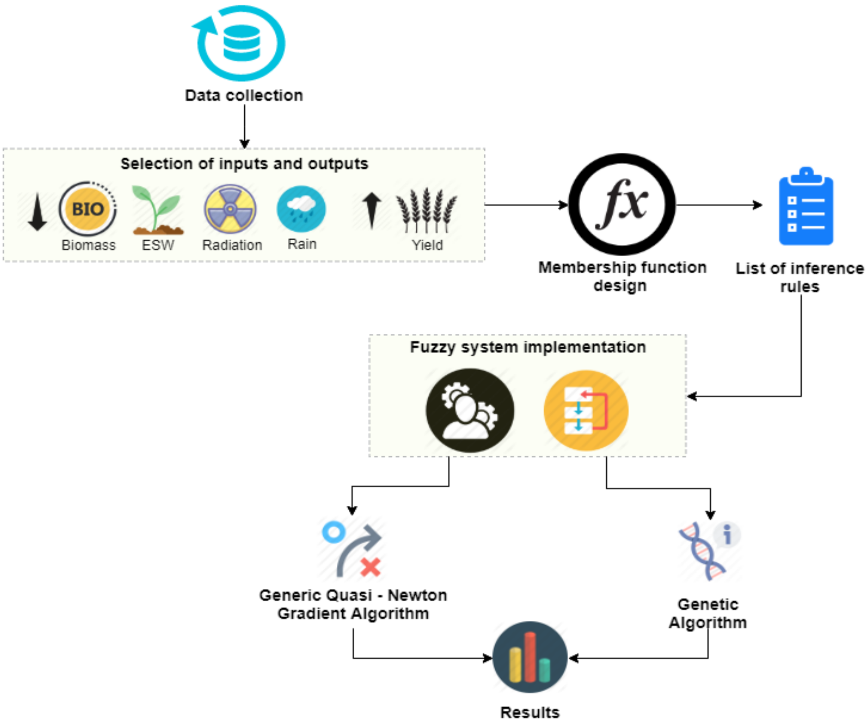

Figure 3. Process flow for fuzzy systems using optimization algorithms

4. OUTCOMES AND DISCUSSION

This section presents the results obtained from the implementation of the four proposed configurations for the fuzzy sets and their respective optimization techniques. Tables describe the performance indices in terms of accuracy and error. On the other hand, the figures represent the temporal behavior of the forecasting made by the best models.

4.1. Generic Quasi-Newton gradient algorithm

Configuration 1: Mamdani system (trimf function) 16 rules.

Configuration 2: Sugeno system (gaussmf function) 10 rules.

Table 1 shows the performance index related to the two configurations for a gradient-based method: error on the first configuration was 0.005433 and 0.006365 on the second. The accuracy percentage was 99.926 and 99.920 accordingly; as the first configuration has the better accuracy and lesser error is declared the best performance outcome. Figures 4 and 5 show the MSE in the first 50 dataset for both configurations. Figures 6 and 7 show the comparison between the real and estimated data for both configurations in a range of 50 datasets. This would allow the researches to exemplify the behavior of the series.

Table 1. Performance indices for configurations applied to the gradient base method

Config	Minimum error	Maximum error	Mean Squared Error (MSE)	Standard deviation (STD)	Root Mean Squared Error (RMSE)	Performance percentage (%)
1	0,000681	0,254324	0,005434	0,442794	0,073712	99,926
2	0,000284	0,447477	0,006365	0,439664	0,079781	99,920

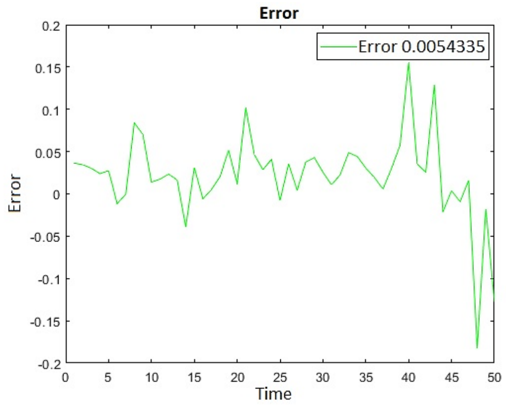

Figure 4. MSE error corresponding to the first configuration for the Quasi-Newton gradient technique

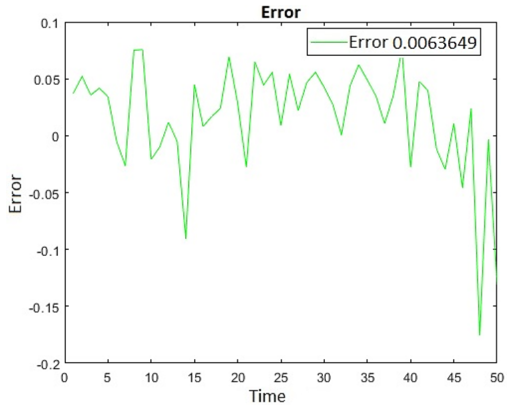

Figure 5. MSE error corresponding to the second configuration for the Quasi-Newton gradient technique

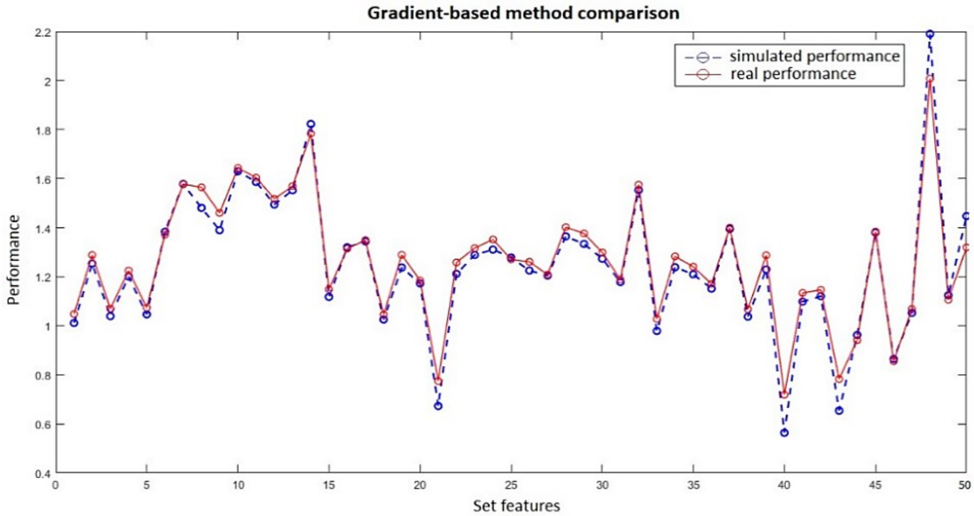

Figure 6. Comparison of real values with those predicted from the first configuration for the Quasi-Newton gradient technique

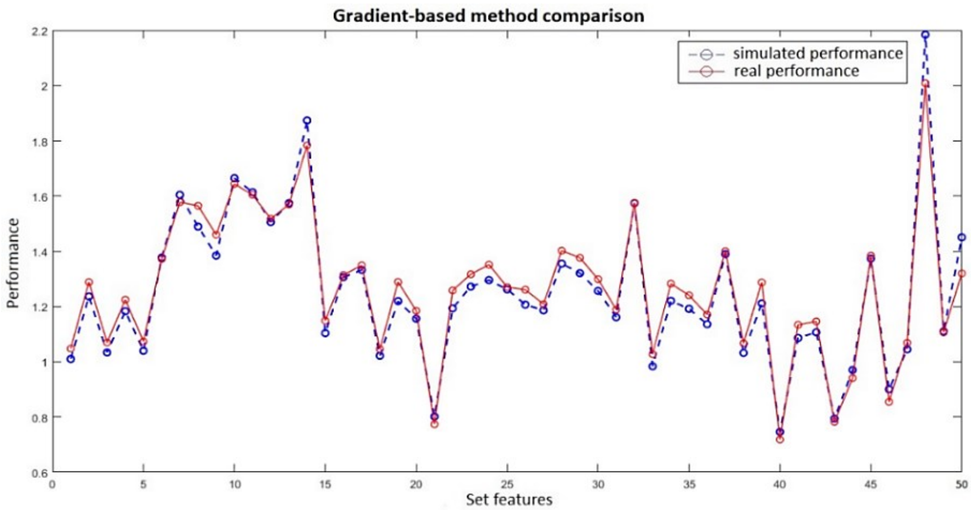

Figure 7. Comparison of real values with those predicted from the second configuration for the Quasi-Newton gradient technique

4.2. Genetic algorithm

Configuration 1: Mamdani system (trimf function) 16 rules.

Configuration 2: Sugeno system (gaussmf function) 10 rules.

Table 2 shows the performance index related to the two configurations for a GA method: error on the first configuration was 0.015139 and 0.005216 on the second. The accuracy percentage was 99.877 and 99.928 accordingly.10 iterations were made in each scenario. The results imply that the second configuration is the one with the lesser error value. Also, the greatest accuracy value on the first configuration is never greater than the lower accuracy value on the second one; this occurs likewise on the MSE and RMSE values. Therefore, as the second configuration is the one with the better accuracy and lesser error, it is declared the best outcome.

Figures 8 and 9 show the MSE in the first 50 dataset for both configurations. Figures 10 and 11 show the comparison between the real and estimated data for both configurations in a range of 50 datasets. Table 3 shows the best values obtained through both configurations. When compared this best values obtained to fuzzy models and ANFIS; MLR and ANN; and SVMs [1, 25, 26] the used methods (combined with an optimization technique) show similar or more accurate outcomes.

Table 2. Performance indices for configurations applied to the method of genetic algorithms

Config	#	Minimum Error	Maximum error	Mean Squared Error (MSE)	Standard deviation (STD)	Root Mean Squared Error (RMSE)	Performance percentage (%)
1	1	0,000284	1,366100	0,022797	0,427430	0,150987	99,849
	2	0,000278	1,366100	0,036932	0,357980	0,192178	99,808
	3	0,001158	1,366100	0,023437	0,422067	0,153091	99,847
	4	0,000205	1,366100	0,018454	0,435044	0,135847	99,864
	5	0,001818	1,366100	0,043785	0,356455	0,209248	99,791
	6	0,000577	1,366100	0,027878	0,417412	0,166967	99,833
	7	0,000496	1,366100	0,020045	0,385376	0,141580	99,858
	8	0,001740	1,366100	0,019911	0,402591	0,141105	99,859
	9	0,000206	1,366100	0,021751	0,419310	0,147482	99,853
	10	0,000165	1,366100	0,015139	0,402565	0,123039	99,877
2	1	0,000027	0,616483	0,005295	0,430868	0,072770	99,927
	2	0,000943	1,049269	0,012609	0,409509	0,112289	99,888
	3	0,001675	1,255310	0,014424	0,407473	0,120102	99,880
	4	0,000346	0,452209	0,005216	0,442515	0,072224	99,928
	5	0,000152	0,280286	0,007739	0,457475	0,087969	99,912
	6	0,000153	0,772559	0,006130	0,428681	0,078293	99,922
	7	0,000111	0,573503	0,009309	0,443502	0,096483	99,904
	8	0,000072	0,640330	0,009869	0,455229	0,099341	99,901
	9	0,000050	0,525823	0,006461	0,459987	0,080383	99,920
	10	0,000027	1,160378	0,012500	0,413769	0,111805	99,888

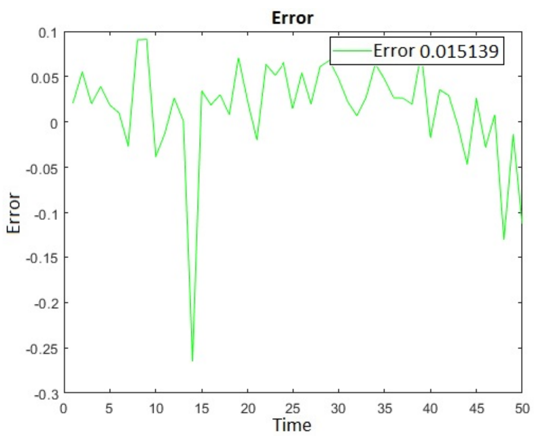

Figure 8. MSE error of the first configuration for the genetic algorithm method

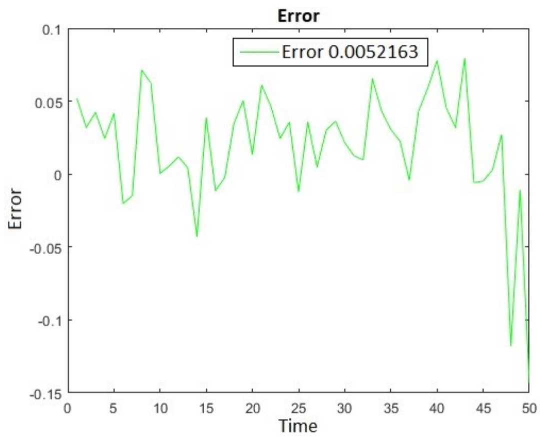

Figure 9. MSE error of the second configuration for the genetic algorithm method

Table 3. Summary table of the best values obtained

Final comparison	Genetic		Gradient	
	MSE	RMSE	MSE	RMSE
Mamdani - 16 rules	0,015139	0,123039	0,005434	0,073712
Sugeno - 10 rules	0,005216	0,072224	0,006365	0,079781

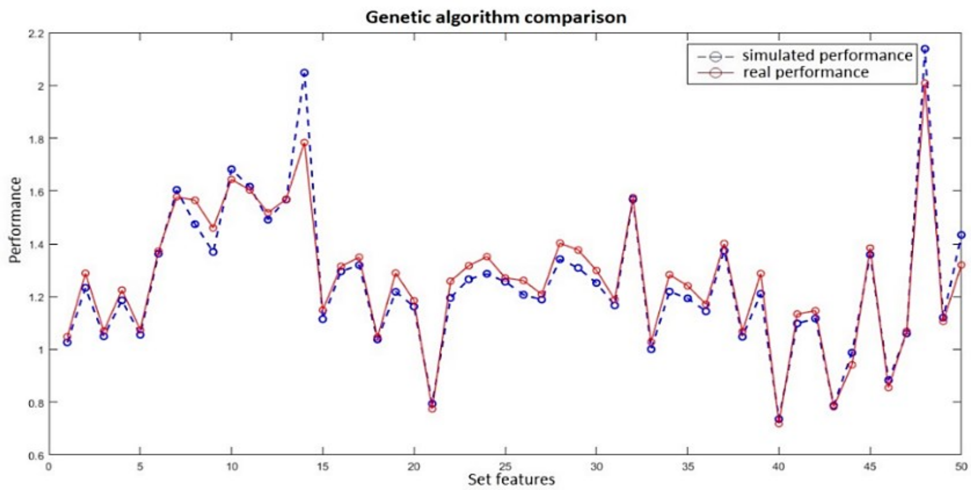

Figure 10. Comparison of real values with those predicted from the first configuration for the genetic algorithm method

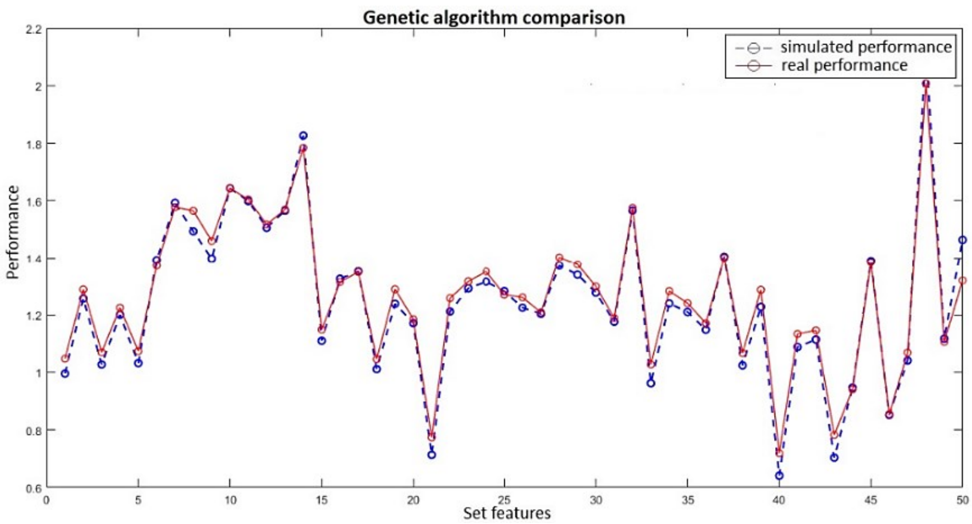

Figure 11. Comparison of real values with those predicted from the second configuration for the genetic algorithm method

5. CONCLUSION

The optimization algorithms used in this research offer an improvement on accuracy for fuzzy sets prediction models and guarantee a great degree of comprehensibility. Nevertheless, an adequate setting on aspects as iteration quantity and optimization time is needed to accomplish the best possible outcome while staying within the average computational resource demand. Performance and error values for the configurations Mamdani (gradient-based algorithm) and Takagi-Sugeno (GA) are relatively close, differentiated only on the fourth decimal for MSE; likewise do RMSE and accuracy. However, it can be concluded which one is the best for wheat crop performance prediction.

The Takagi-Sugeno configuration (GA, 10 rules) showed the best performance for the given quantitative data: its outcomes were 99.928% on performance, 0.005216 on MSE, and 0.072224 on RMSE. Such values are the lowest obtained in the process. Regarding the use of fuzzy sets, the Takagi-Sugeno system reacts more accurately to a GA, and the Mamdani system to a gradient-based algorithm. This would give an insight on how fuzzy sets respond to different optimization techniques, both deterministic or meta-heuristic. The use of these techniques for the prediction of wheat crop performance is an advisable alternative to improve the performance of the agricultural field.

REFERENCES

[1] A. Shastry, Sanjay H. A., and M. Hegde, "A parameter based ANFIS model for crop yield prediction," *2015 IEEE International Advance Computing Conference (IACC)*, pp. 253-257, 2015.

[2] A. H. Mohammed, A. M. Gad Allah, and H. A. Hefny, "Fuzzy time series approach for optimizing crops planting dates with climate changes," *2014 10th International Computer Engineering Conference (ICENCO)*, pp. 36-41, 2014

[3] D. K. Gupta, *et al.*, "Crop variables estimation by adaptive neuro-fuzzy inference system using bistatic scatterometer data," *2015 International Conference on Microwave and Photonics (ICMAP)*, pp. 1-2, 2015.

[4] F. Jawad, *et al.*, "Analysis of Optimum Crop Cultivation using Fuzzy System," *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, pp. 1-6, 2016.

[5] S. N. Deve and P. Geetha, "Land use modeling for sustainable crop development using fuzzy approach," *2017 International Conference on Communication and Signal Processing (ICCSP)*, pp. 1473-1476, 2017.

[6] S. B. Zinnat and D. M. Abdullah, "Design of a fuzzy logic based automated shading and irrigation system," *2014 17th International Conference on Computer and Information Technology (ICCIT)*, pp. 170-173, 2014

[7] A. Daniel, P. Sharma, and R. Srivastava, "Fuzzy Based Prediction Model Using Rainfall Parameter for North eat India Maize production," *2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, pp. 1-6, 2018.

[8] A. Anitha and D. P. Acharjya, "Crop suitability prediction in Vellore District using rough set on fuzzy approximation space and neural network," *Neural Comput. Appl.*, vol. 30, no. 12, pp. 3633-3650, 2018.

[9] G. Prabakaran, D. Vaithiyanathan, and M. Ganesan, "Fuzzy decision support system for improving the crop productivity and efficient use of fertilizers," *Comput. Electron. Agric.*, vol. 150, no. March, pp. 88-97, 2018.

[10] I. C. Valenzuela, R. G. Baldovino, A. A. Bandala, and E. P. Dadios, "Optimization of Photosynthetic Rate Parameters using Adaptive Neuro-Fuzzy Inference System (ANFIS)," *2017 International Conference on Computer and Applications (ICCA)*, pp. 129-134, 2017.

[11] J. R. Pérez López, "Contribución a los métodos de optimización basados en procesos naturales y su aplicación a la medida de antenas en campo próximo," Universidad de Cantabria, 2007.

[12] A. Darwish, "Bio-inspired computing: Algorithms review, deep analysis, and the scope of applications," *Futur. Comput. Informatics J.*, vol. 3, no. 2, pp. 231-246, 2018.

[13] X. S. Yang, "Introduction to algorithms for data mining and machine learning," 1st ed. London, *United Kingdom: Elsevier Inc*, 2019.

[14] R. Hauser, "Quasi-Newton Methods," Oxford, 2006. [Online]. Available: http://www.numerical.rl.ac.uk/people/nimg/oupartc/lectures/raphael/lectures/lec4slides.pdf.

[15] J. G. Marín Morales, "Factores que afectan el crecimiento de las plantas," *Corporación Colomb. Investig. Agropecu.*, 2018, Accessed: Apr. 08, 2020. [Online]. Available: http://hdl.handle.net/20.500.12324/22033.

[16] A. Martínez Romero and A. Leyva Galán, "La biomasa de los cultivos en el agroecosistema. Sus beneficios agroecológicos," *Cultiv. Trop.*, vol. 35, no. 1, pp. 11-20, 2014,

[17] N. Dalgliesh, *et al.*, "Expanding the area for Rabi-season cropping in southern Bangladesh," 1st ed. *Australia: ACIAR,* 2012.

[18] J. West, "Multi-criteria evolutionary algorithm optimization for horticulture crop management," *Agric. Syst.*, vol. 173, no. January, pp. 469-481, 2019.

[19] C. Dai, M. Yao, Z. Xie, C. Chen, and J. Liu, "Parameter optimization for growth model of greenhouse crop using genetic algorithms," *Appl. Soft Comput. J.*, vol. 9, no. 1, pp. 13-19, 2009.

[20] D. C. H. Nguyen, J. C. Ascough, H. R. Maier, G. C. Dandy, and A. A. Andales, "Optimization of irrigation scheduling using ant colony algorithms and an advanced cropping system model," *Environ. Model. Softw.*, vol. 97, pp. 32-45, 2017.

[21] P. Ariyatanakatawong and S. Rimcharoen, "Enhanced Genetic Algorithm with Risk Management Model for Optimizing Rice Plantation," *2018 18th International Symposium on Communications and Information Technologies (ISCIT)*, pp. 509-513, 2018.

[22] I. Esfandiarpour-Boroujeni, E. Karimi, H. Shirani, M. Esmaeilizadeh, and Z. Mosleh, "Yield prediction of apricot using a hybrid particle swarm optimization-imperialist competitive algorithm- support vector regression (PSO-ICA-SVR) method," *Sci. Hortic. (Amsterdam)*, vol. 257, 2019.

[23] X. Jin, Z. Li, H. Feng, Z. Ren, and S. Li, "Estimation of maize yield by assimilating biomass and canopy cover derived from hyperspectral data into the AquaCrop model," *Agric. Water Manag.*, vol. 227, 2020.

[24] F. A. Coutelieris and A. Kanavouras, "Experimentation Methodology for Engineers," 1st ed. *Springer International Publishing*, 2018.

[25] P. S. Maya Gopal and R. Bhargavi, "A novel approach for efficient crop yield prediction," *Comput. Electron. Agric.*, vol. 165, 2019.

[26] N. Gandhi, O. Petkar, L. J. Armstrong, and A. K. Tripathy, "Rice crop yield prediction in India using support vector machines," *2016 13th International Joint Conference on Computer Science and Software Engineering, JCSSE 2016*, pp. 1-5, 2016.