

Ahmed A. Ewees* and Mohamed Abd Elaziz

Improved Adaptive Neuro-Fuzzy Inference System Using Gray Wolf Optimization: A Case Study in Predicting Biochar Yield

<https://doi.org/10.1515/jisys-2017-0641>

Received December 28, 2017; previously published online September 25, 2018.

Abstract: This paper presents an alternative method for predicting biochar yields from biomass thermochemical processes. As biochar is considered a renewable and sustainable energy source, it has received more attention. Several methods have been presented to predict biochar, such as neural network (NN) and least square support vector machine (LS-SVM). However, each of them has its own drawbacks, such as getting stuck in a local optimum, which occurs in NN, and lack of uncertainty and time complexity, as in LS-SVM. Therefore, this paper avoids this limitation by using a hybrid method between the adaptive neuro-fuzzy inference system (ANFIS) and gray wolf optimization (GWO) algorithm. The proposed method is called ANFIS-GWO, which consists of two stages. In the first stage, GWO is used to learn the parameters of ANFIS using the training set. Meanwhile, in the second stage, the testing set is used to evaluate the performance of the proposed ANFIS-GWO method. Three experiments were performed to assess the performance of the proposed method. The first experiment used a set of UCI (University of California, Irvine) benchmark datasets to evaluate the effectiveness of ANFIS-GWO. The aim of the second experiment was to evaluate the performance of the proposed ANFIS-GWO method to predict biochar yield from manure pyrolysis. The third experiment aimed to estimate the values of input parameters of pyrolysis that maximize biochar production. The obtained results were compared to those of other methods, such as ANFIS using gradient descent, practical swarm optimization, genetic algorithm, whale optimization algorithm, sine-cosine algorithm, and LS-SVM. The results of the ANFIS-GWO method were >35% of the standard ANFIS and also better than those of other methods.

Keywords: Gray wolf optimization (GWO), adaptive neuro-fuzzy inference system (ANFIS), renewable energy production, biochar prediction.

1 Introduction

The world's energy use has increased >10-fold during the 20th century. Energy purchases account for 5–10% of the entire national product in developed economies. However, in some of those countries, imports of fossil fuel may cost over half the value of overall exports. These economic realities are unsustainable, and pose a challenge to sustainable development [50]. Furthermore, fossil fuel is decreasing dramatically; therefore, the use of renewable energy from various sources has significantly grown in recent years. The renewable energy sources include biomass, solar radiation, ocean waves, wind, geothermal heat, and others.

Biochar is considered an attractive renewable energy source [51]. It is created from biomass through thermochemical methods, the yields of which rely on the methodological circumstances. Pyrolysis is the most popular process used to produce biochar [22]. Based on residence time and heating rate, pyrolysis is classified into two kinds: slow and fast. Biochar is created from biomass by using slow pyrolysis through a comparatively long residence period and a low heating degree, whereas fast pyrolysis uses a high heating degree and little residence period to yield a biochar.

*Corresponding author: **Ahmed A. Ewees**, University of Bisha, Bisha, Kingdom of Saudi Arabia; and Department of Computer, Damietta University, Damietta, Egypt, e-mail: eweas@du.edu.eg

Mohamed Abd Elaziz: Department of Mathematics, Faculty of Science, Zagazig University, Zagazig, Egypt

The most attractive feature of biochar is its ability to represent a sustainable, inexpensive, and easy-to-produce method. Also, biochar has many applications in the economic-, technical-, soil-, and climate-related aspects [4, 21, 26, 31, 47]. These applications include oil amendment, energy, gas storage, water purification, and catalysis. Biochar can be used as a precursor for creating a catalyst for syngas made from the gasification of biomass, which includes a substantial quantity of tar [9]. Increase in agricultural productivity by using biochar for soil improvement can be translated into an increase in soil fecundity [40].

The prediction of biochar may be considered an optimization problem, where various artificial intelligence models are implemented to predict sustainable and renewable energy, such as solar energy [33, 37], wind energy [10], hydropower [32], bioenergy [39], geothermal energy [49], and hybrid systems [24].

Several artificial intelligence algorithms have been applied in many renewable energy applications, such as the artificial NN (ANN) [27, 46]. However, ANN has some drawbacks, such as it may get stuck in a local point; thus, the least square support vector machine (LS-SVM) is used as a method to resolve the drawbacks of ANN [8]. The performance of the LS-SVM model in predicting biochar was higher than the performance of the ANN model. Cao et al. [8] gave two reasons for this result: (i) LS-SVM uses the structural risk minimization principle that works to minimize an upper limit for the error of generalization instead of minimizing the error of training. In the other words, ANN uses the empirical risk minimization principle. (ii) LS-SVM can achieve a globally optimal solution, but ANN could not in most cases.

Nevertheless, LS-SVM suffers from some drawbacks, such as its performance is highly dependent on the selected kernel function. Also, it is relatively slow because the parameters of SVM must be tuned by conducting a coarse grid search and then a fine grid search. It cannot use the fuzzy rules as the change rules between the biochar characteristics and operating conditions occur in a fuzzy manner. Therefore, the fuzzy set theory is applied to address these cases [29]. Therefore, fuzzy sets have many interests in renewable energy issues [48], and the fuzzy inference system was used widely in renewable energy. A fuzzy-based model uses logical operators and IF-THEN rules in order to build qualitative relationships between its variables. The neuro-fuzzy systems combine the advantages of each of fuzzy logic and ANN in one model, through applying the ANN in adapting rule-based fuzzy systems. A particular model in neuro-fuzzy systems is the adaptive neuro-fuzzy inference system (ANFIS), which has achieved great outcomes in modeling non-linear functions. The membership function parameters of ANFIS are obtained from a dataset that defines the system operation. The ANFIS uses the features of the dataset to learn, and modifies the system parameters based on an error criterion [23]. The ANFIS has recorded many successful implementations in predicting the results of renewable energy, such as in solar energy [41], wind energy [1, 3], and bioenergy [13, 18].

However, the ANFIS model has some drawbacks when it is used for real applications; these drawbacks result from training the parameters of the memberships function and weights between layers of the ANFIS model. The gradient descent approaches are popular algorithms that are used to learn the parameters of ANFIS. However, in each repetition, the gradient is computed and its performance is affected by the initial point. Also, it can get stuck in the local point and therefore is not a global solution for parameters that can be determined [23]. To solve these drawbacks, meta-heuristics like genetic algorithm (GA) [30] and particle swarm optimization (PSO) [16, 45] were applied. However, GAs have always been criticized for their slow convergence speed, whereas PSO could encounter premature convergence at the later stage of the search process and is sensitive to neighborhood topology. Thus, the gray wolf optimization (GWO) algorithm is used to deal with this issue [36].

The GWO is a new meta-heuristic that emulates the behavior of gray wolves in nature [36]. In GWO, four kinds of wolves (i.e. α , β , δ , and ω) are utilized, where α wolves control the β , which in turn control the δ , whereas all the higher wolves control the ω wolves. There are three steps used to mimic the hunting behaviors (i.e. tracking, encircling, and attacking the prey). The GWO algorithm has a small number of parameters compared to other algorithms, and its performance has been established in several applications, such to enhance the performance of an NN, as in Refs. [17, 38]. Also, Robandi [42] used GWO to estimate photovoltaic parameters. It is applied to solve the system reliability optimization as in Ref. [25], which provides better results than other algorithms. In addition, it has been applied in optimizing support vector machines and clustering applications [6] (for more applications, see Ref. [20]). From the promised results of GWO in these applications, we decide to use it in this paper.

The main contributions of the current work are to (i) improve the performance of ANFIS through learning its parameters using the GWO algorithm, and (ii) to apply the improved ANFIS model as a new prediction method for renewable energy production, especially biochar production. Therefore, the proposed method for predicting biochar, which is called ANFIS-GWO, consists of two stages: (i) training ANFIS and (ii) testing. In the first stage, the ANFIS parameters are estimated using the GWO algorithm in which each solution of the population of GWO contains the parameter's value. Meanwhile, the ANFIS-GWO algorithm is evaluated in the second phase by using the testing set (which is not used before in the training stage) as input and then predicting the output and computing the performance.

This paper is arranged as follows: in Section 2, the standard ANFIS model, GWO, and fuzzy C-means (FCM) are explained. In Section 3, the proposed method is introduced. Section 4 gives the experimental results and discussion. The conclusion is given in Section 5.

2 Preliminaries

2.1 ANFIS

Jang [23] introduced ANFIS, which is a hybrid model that combines fuzzy logic and NNs. ANFIS is based on the Takagi-Sugeno inference model that creates a non-linear mapping from input to the output space by using fuzzy IF-THEN rules. The ANFIS model contains five layers, as shown in Figure 1.

The crisp inputs x and y to the node of the first layer and the output O_{1i} of this node are defined as

$$O_{1i} = \mu_{A_i}(x), \quad i = 1, 2, \quad O_{1i} = \mu_{B_{i-2}}(y), \quad i = 3, 4, \quad (1)$$

where A_i and B_i are the membership values of the generalized Gaussian membership function defined as [23]

$$\mu(x) = e^{-\left(x - \frac{p_i}{\sigma_i}\right)^2}, \quad (2)$$

where p_i and σ_i are the premise parameters. In the second layer, the node's output is the firing strength of a rule, as

$$O_{2i} = \mu_{A_i}(x) \times \mu_{B_{i-2}}(y). \quad (3)$$

The node's output in the third layer is the normalized firing strength as

$$O_{3i} = \bar{w}_i = \frac{\omega_i}{\sum_{i=1}^2 \omega_i}. \quad (4)$$

The node in layer 4 is an adaptive node, and its output is computed as

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i), \quad (5)$$

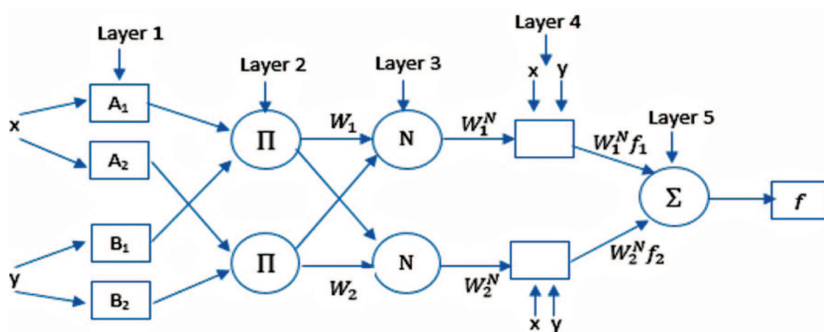


Figure 1: Layers of the ANFIS Model.

where p_i , q_i , and r_i are the consequent parameters of the node i . In the last layer, there exists only one node whose output is computed by using the following equation:

$$O_5 = \sum_i \bar{w}_i f_i. \quad (6)$$

There are two sets of adjustable parameters of the ANFIS model: the premise and the consequent parameters. The least square method (LSM) can be used to determine the fitness values of the consequent parameters. However, if the premise parameters are not steady, the search space will be wider and the convergence of training will be slower. Therefore, the hybrid learning techniques can be used to overcome this problem. One of these hybrid learning algorithms is the hybrid between the back-propagation algorithm and the LSM. This algorithm involves a two-step process [23].

In the first one, if the premise parameters are steady, the functional signals will propagate to layer 4, where the LSM specifies the consequent parameters. Thereafter, the consequent parameters will be kept fixed.

The adaptation methods of most fuzzy inference systems rely on the back-propagation algorithm that is applied to deal with parameter optimization in general. This traditional optimization technique can get trapped in a local optimum. To fix this problem, evolutionary methods like GA have been widely applied. Nevertheless, these methods could not achieve the promised results in all experimental cases and need much computation time; therefore, we use the GWO algorithm to determine the optimal weights of ANFIS and reduce the time complexity.

2.2 GWO Algorithm

The GWO algorithm is a swarm algorithm that emulates the social hierarchy of wolves [36], wherein wolves aim to determine the location of prey. The hierarchy of wolves consists of three layers: the first is the α wolves that correspond to the fittest wolves; meanwhile, the other two layers are called β and δ wolves, which represent the second and third best wolves, respectively. Therefore, these wolves are responsible for searching for prey in the search space, while the other wolves follow them. Three steps are used to simulate the hunting behavior: (i) encircling, (ii) tracking, and (iii) attacking the prey. The definition of encircling is given in Eq. (7) [36]:

$$D(t+1) = |A \cdot X_p(t+1) - X(t+1)|, \quad (7)$$

$$X(t+1) = |X_p(t+1) - F \cdot D(t+1)|, \quad (8)$$

where D is the distance between the wolf X and the prey X_p , while F and A represent the coefficient vectors that are computed as follows:

$$F = 2f \cdot r_1 - f, \quad (9)$$

$$A = 2r_2, \quad (10)$$

where r_1 and r_2 represent random vectors that belong to the interval $[0, 1]$, and the value of parameter a is decreased from 2–0 in linear form with each iteration increased.

The position of any wolf X in the current population can be updated, as in Figure 2, according to the position of α , β , and δ as

$$X = \frac{(X_1 + X_2 + X_3)}{3}, \quad (11)$$

where

$$X_1 = |X_\alpha - A_\alpha \cdot D_\alpha|, \quad X_2 = |X_\beta - A_\beta \cdot D_\beta|, \quad X_3 = |X_\delta - A_\delta \cdot D_\delta|, \quad (12)$$

$$D_\delta = |C_3 \cdot X_\delta - X|, \quad D_\alpha = |C_2 \cdot X_\alpha - X|, \quad D_\beta = |C_2 \cdot X_\beta - X|. \quad (13)$$

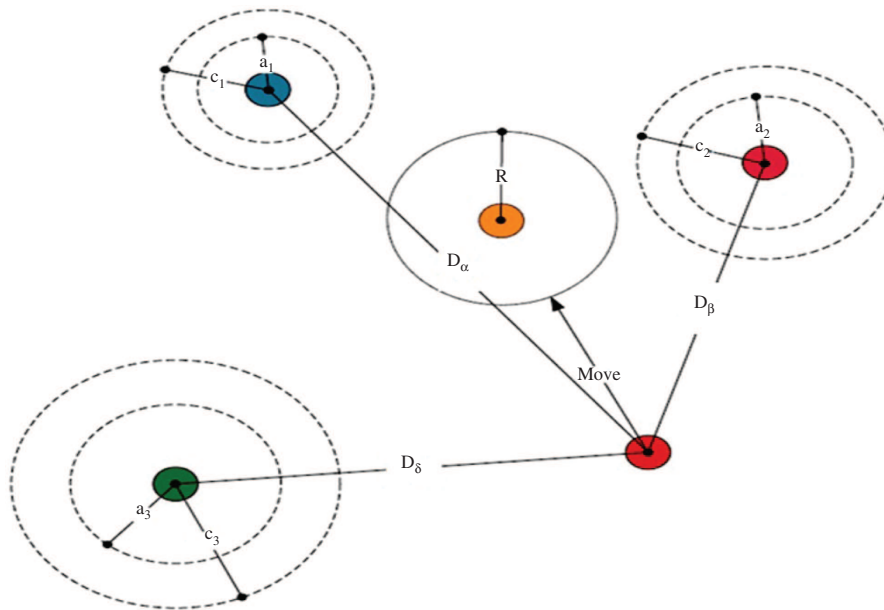


Figure 2: Updating the Wolves' Positions in the GWO Algorithm.

2.3 FCM Clustering Method

The most important concept in ANFIS is defining the number of membership functions. This can be considered a clustering problem; therefore, the FCM is used to achieve a small number of fuzzy rules. FCM is a clustering method; it was improved by Bezdek et al. [7], who determined the degree of data belonging to different clusters through minimizing the following objective function:

$$T_r = \sum_{(i=1)}^N \sum_{(t=1)}^C g_{it}^r \|x_i - c_t\|^2, \quad 1 \leq r \leq \infty, \quad (14)$$

where $r > 1$ is a real number, g_{it} represents the degree of membership of the measured data $x_i \in R^d$ to belong to the cluster of center $c_t \in R^d$. Fuzzy partitioning is executed through minimizing Eq. (14), with updating of the membership (g_{it}) and the center of clusters (c_t) using Eqs. (15) and (16):

$$g_{it} = \frac{1}{\sum_{(k=1)}^C \left(\frac{\|x_i - c_t\|}{\|x_i - c_k\|} \right)^{\left(\frac{2}{r-1} \right)}}. \quad (15)$$

$$c_t = \frac{\sum_{(i=1)}^N g_{it}^r x_i}{\sum_{(i=1)}^N g_{it}^r}. \quad (16)$$

This iteration will stop when $\max_{it} \{ |g_{it}^{(k+1)} - g_{it}^{(k)}| \} < \epsilon$, where $\epsilon \in [0, 1]$ is an ending criterion. The previous steps are repeated until the stopping conditions are reached.

3 Proposed Method

This section introduces the proposed method for predicting the biochar yield from manure pyrolysis. This model is based on a hybrid between the GWO and ANFIS. It is called ANFIS-GWO, in which the parameters of the ANFIS are determined by using the GWO algorithm. This model has five layers similar two traditional ANFIS. The nodes in layer 1 are represented by input variables (heating rate, pyrolysis temperature, holding

time, moisture content, sample mass). The nodes in layer 2 are the membership functions of the input variables. The fuzzy logic rules are represented by the nodes in layer 3. The nodes of layer 4 use the consequent part of the Takagi-Sugeno-Kang model. The output of layer 5 is the biochar. In the learning stage, the GWO is used to determine the best value of the weights between layers 4 and 5, as well as to train the membership function according to the input variable.

The proposed method starts by normalizing the dataset and dividing it into two groups (training and testing). Then, the FCM is used to define the number of clustering (i.e. the number of membership functions). The next step is constructing the ANFIS using the FCM output. The parameters in ANFIS are updated based on the GWO algorithm, where the GWO explores different regions of the search space that has many local minima and then reduces the domain of search to the area that contains the global solution. In the training phase, the error information between the actual output and the corresponding predicted values are used to update the parameters (weights between layers 4 and 5 and parameters of Gaussian membership function). Where the GWO starts by generating a population X , the population with a random position for each wolf the size of X is set to N and dimension D , which represents the number of ANFIS parameters. Then, the objective function for a solution inside the population X is computed and the value of α , β , and γ are updated based on the smallest three fitness functions. The objective function (obj) in our model is defined as

$$obj = \|y - \hat{y}\|^2, \quad (17)$$

where y and $\hat{y}$ are the original biochar value and its predicted value, respectively. This function represents the summation of the square error between the predicted biochar value $\hat{y}$ and the original value (y). Therefore, the best solution is that which has the minimum objective function value.

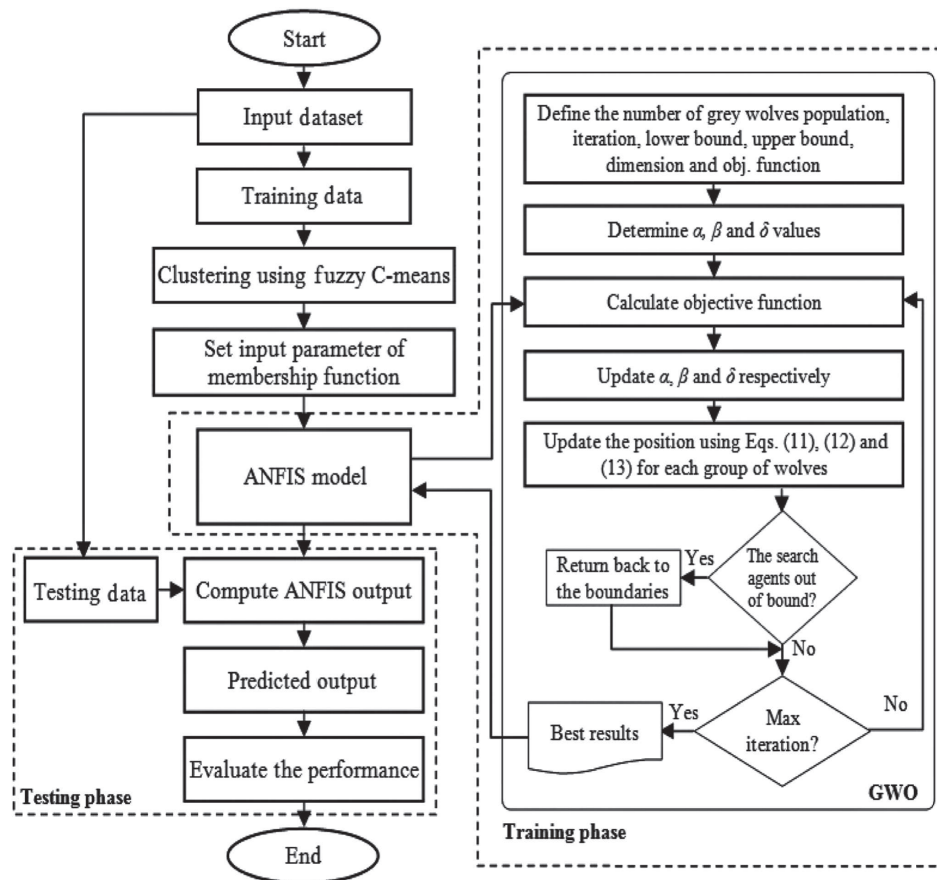


Figure 3: Proposed ANFIS-GWO Method.

The next step in the GWO algorithm is to update the position of other solutions according to the value of α , β , and γ using Eqs. (11)–(13). These steps previously performed are still repeated until the stop condition is satisfied. Then, the best solution is passed to the ANFIS model. The training phase is completed if the stop conditions (maximum number of iterations and error less than the small value) are satisfied. Then, the ANFIS is constructed based on the parameters coming from the best solution. In the testing phase, the test data set is applied to the proposed method that predicts the output (biochar) based on input parameters and evaluates the performance of the model. The proposed method is explained in Figure 3.

4 Experimental Results and Discussion

In this study, three experiments were performed to test the performance of ANFIS-GWO. The first one was applied to evaluate the performance of the proposed method in achieving a minimum error using six regression datasets that are collected from the UCI (University of California, Irvine) repository [28]. The results of this experiment are recorded in Section 4.3. The second experiment was applied to predict a biochar yield from manure pyrolysis, and it is explained in Section 4.4. The last one was used to estimate the values of input parameters of pyrolysis that maximize the biochar production, as in Section 4.5.

4.1 Performance Measures

To assess the performance of the proposed ANFIS-GWO method and to test the quality of solutions, several measures were used, which are defined as follows:

1. Mean square error (MSE): measures the dispersion of the data around zero. It is computed as

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (18)$$

where $\hat{y}_i$ is the i^{th} predicted element, y_i is the i^{th} measured element, and n is the number of samples. y_i is the average of the corresponding predicted value.

2. Mean absolute error (MAE): measures the mean absolute deviation of the values of the output from the values of the target. It is computed as

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (19)$$

where $\hat{y}_i$ is the i^{th} predicted element, y_i is the i^{th} measured element, and n is the number of samples. y_i is the average of the corresponding predicted value.

3. Root MSE (RMSE): the square of MSE to measure the dispersion of the data around zero, as in the following equation:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (20)$$

4. Average absolute percent relative error (AAPRE): computes the relative absolute deviation of the predicted output from the experimental output; the lowest AAPRE value indicates the most accurate prediction. It is defined mathematically as follows:

$$AAPRE = \frac{100}{n} \left(\sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \right). \quad (21)$$

5. Coefficient of determination (R^2): measures the relationship between the experimental data and the predicted data, and the algorithm that has R^2 value closer to 1. R^2 is defined as

$$R^2 = 1 - \sum_{i=1}^n \frac{(y_i - \hat{y}_i)}{(y_i - \bar{y}_i)}. \quad (22)$$

The highest R^2 value indicates the most accurate prediction.

6. Standard deviation (SD):

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}_i)^2}. \quad (23)$$

4.2 Experimental Settings

In all experiments, a 10-fold cross-validation method was used to divide the data into training and testing sets. The 10-fold cross-validation split the data into 10 groups, assigning nine groups as the training set and one group as the testing set. These processes are repeated 10 times until all groups become a testing set. The output is computed as the average of the total 10 runs.

The experiments were implemented in Matlab R2014b and MS Windows 10 (64 bit) over Intel Core2Duo processor and 4 MB of RAM. The parameters of all optimization algorithms were set as follows: size of the population (n) = 25, maximum iteration = 100, lower bound = -5 , and upper bound = 5 . These parameters were used for fair comparison between the algorithms. However, we found that in the case of increasing the population size to become >25 , the performance increased little but the algorithms took very large CPU time(s). Also, we noticed that, if the population size decreased to <25 , the performance is decreased (the same for the maximum number of iterations).

4.3 Experiment Series 1: Evaluating the Performance of ANFIS-GWO

In this section, six datasets are used to evaluate the performance of the proposed method (ANFIS-GWO), which are collected from the UCI repository [28]. The descriptions of these datasets are given in Table 1. In the following experiment, we refer to the Yeast, Wine, AirfoilSelfNoise, Forestfires, Heart, and Housing datasets as Dataset1, Dataset2, Dataset3, Dataset4, Dataset5, and Dataset6, respectively. In tables and figures, ANFIS-GWO, ANFIS-PSO, and ANFIS-GA are referred to as GWO, PSO, and GA, respectively. The results of the proposed ANFIS-GWO method are compared with the other three algorithms, namely ANFIS-PSO [5], ANFIS-GA [44], and the original ANFIS.

The experiments were applied 15 times, and the average of all measures are recorded in Table 2 and Figures 4–7.

From Table 2 and Figures 4–7, we can conclude that ANFIS-GWO outperformed the other algorithms in four datasets out of six in terms of MSE and RMSE, whereas it outperformed three datasets out of six in terms of MAE. Meanwhile, ANFIS-PSO came in second place. According to the time complexity, we can see that the

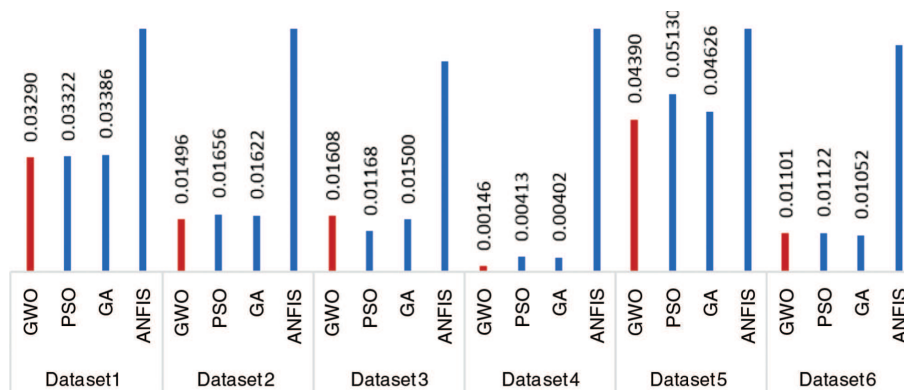
Table 1: Description of the Six Datasets Used to Evaluate the Proposed Method.

Code name	Dataset name	Attribute number	Instance number
Dataset1	Yeast	8	1484
Dataset2	Wine	13	178
Dataset3	AirfoilSelfNoise	6	1503
Dataset4	Forestfires	13	517
Dataset5	Heart	13	303
Dataset6	Housing	13	506

Table 2: Results of ANFIS-GWO Over Six Datasets against Three Algorithms.

Dataset	Algorithm	MSE	RMSE	MAE	Time
Dataset1	GWO	0.03290	0.18106	0.13120	45.54160
	PSO	0.03322	0.18203	0.12499	46.86776
	GA	0.03386	0.18369	0.13146	43.55836
	ANFIS	2.45680	1.56283	1.02184	30.63989
Dataset2	GWO	0.01496	0.12173	0.09816	35.86977
	PSO	0.01656	0.12831	0.10301	36.20687
	GA	0.01622	0.12660	0.10128	36.39580
	ANFIS	0.30172	0.51610	0.32813	11.56986
Dataset3	GWO	0.01608	0.12671	0.09875	31.13066
	PSO	0.01168	0.10800	0.08258	26.91375
	GA	0.01500	0.12238	0.09444	30.07782
	ANFIS	0.06087	0.21176	0.16457	9.97239
Dataset4	GWO	0.00146	0.03332	0.01739	37.00145
	PSO	0.00413	0.06071	0.02144	47.11759
	GA	0.00402	0.05797	0.02055	37.31711
	ANFIS	7.06525	1.94732	0.21341	22.18185
Dataset5	GWO	0.04390	0.20857	0.15655	35.51686
	PSO	0.05130	0.22547	0.16682	37.21948
	GA	0.04626	0.21367	0.15858	36.25798
	ANFIS	6.54830	1.64775	0.56228	19.37667
Dataset6	GWO	0.01101	0.10477	0.07590	46.51049
	PSO	0.01122	0.10501	0.07123	43.33064
	GA	0.01052	0.10221	0.07191	48.34658
	ANFIS	0.06536	0.21647	0.08408	28.60306

Boldface indicates the best result.

**Figure 4:** Results of the MSE Measures.

original ANFIS is faster than all other compared algorithms, although it fails to obtain the smallest error in all datasets and the differences in the time complexity between it and the other algorithms are still small. Thus, we excluded its results from our comparison because although it has the best time complexity, it has the worst performance. According to the time complexity, it can be seen that ANFIS-GWO achieved the lowest time complexity in three datasets out of six, whereas ANFIS-PSO came in second place followed by ANFIS-GA.

To provide more evidence about the quality of the ANFIS-GWO method in prediction, a non-parametric Wilcoxon rank sum test was used. It was applied to the analysis of the median of the ANFIS-GWO with the median of the compared algorithms according to the RMSE measure at a significance level 0.05. The value $p < 0.05$ indicates that there is a statistically significant difference between ANFIS-GWO and the other

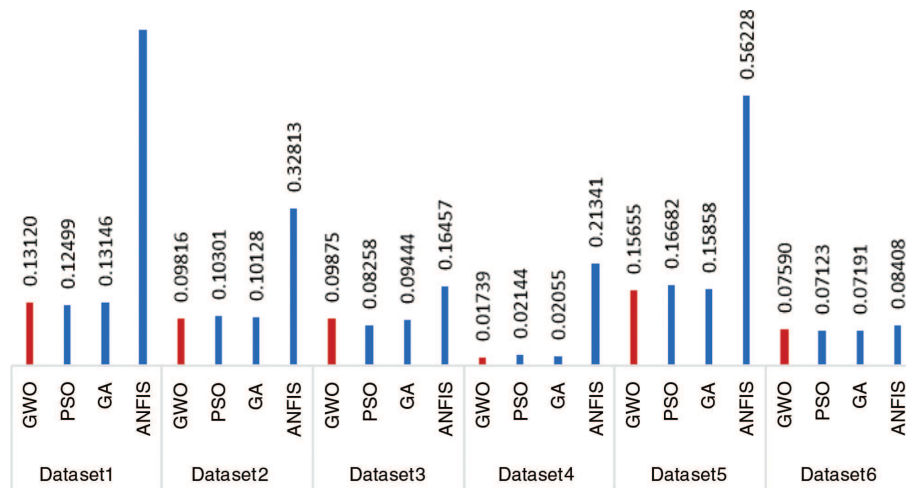


Figure 5: Results of the MAE Measures.

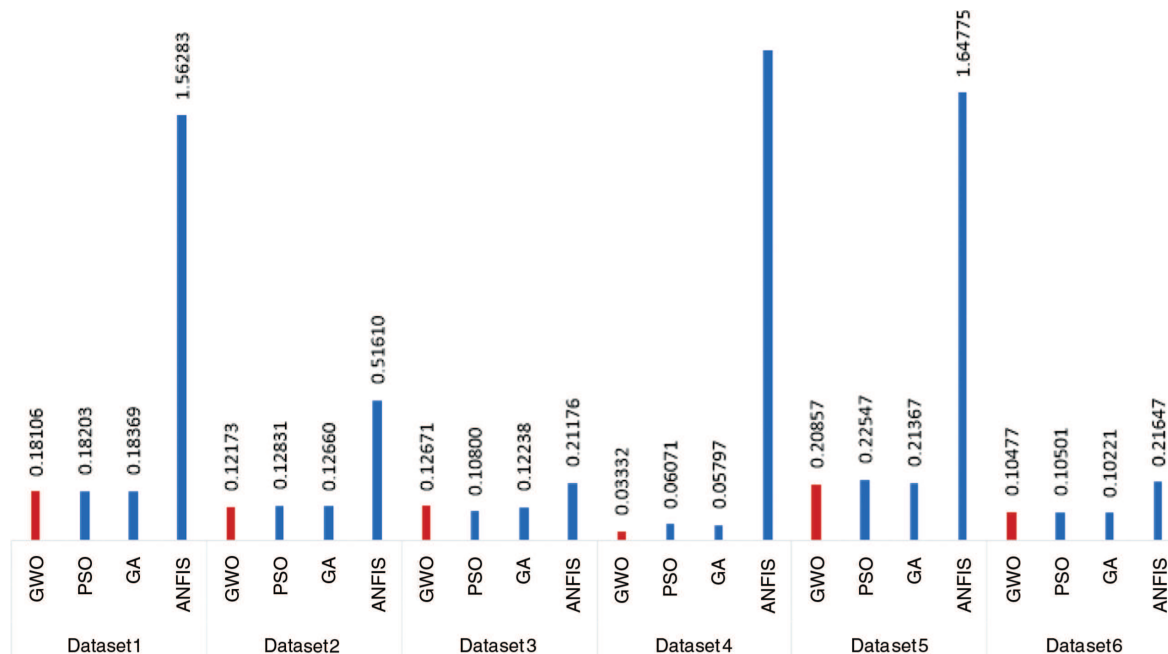


Figure 6: Results of the RMSE Measures.

algorithms. The results of the Wilcoxon test are listed in Table 3, and it can be observed that there are statistically significant differences between ANFIS-GWO and the compared algorithms in general. In this context, there was a statistical difference between ANFIS-GWO and the original ANFIS over all experiments, whereas there were statistical differences between ANFIS-GWO and both ANFIS-PSO and ANFIS-GA in all experiments except for Dataset1 and Dataset6, and except for Dataset5 for ANFIS-GA.

4.4 Experiment Series 2: Predicting Biochar Yield

In this experiment, the proposed algorithm was used to improve the prediction performance of the biochar yield. However, this section starts with the problem formulation, followed by the results of the comparison algorithms.

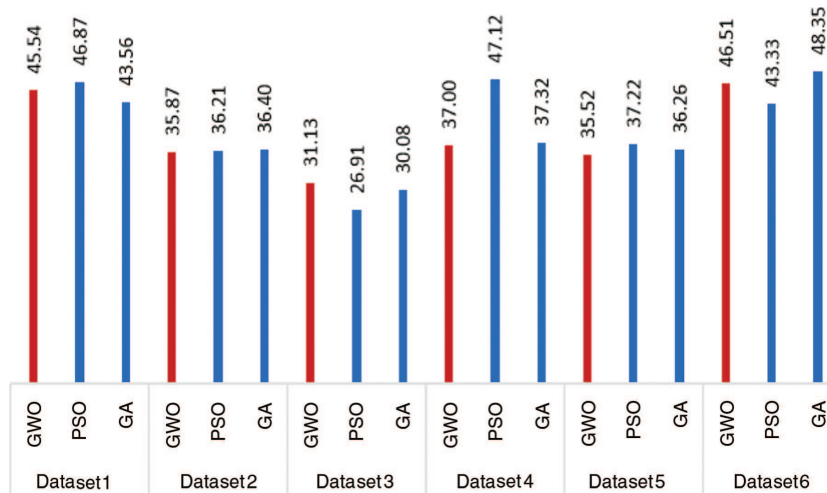


Figure 7: Results of Time Complexity.

Table 3: Results of the Wilcoxon Statistical Test.

Dataset	PSO	GA	ANFIS
Dataset1	0.455	0.263	0.000
Dataset2	0.038	0.042	0.000
Dataset3	0.000	0.038	0.001
Dataset4	0.001	0.003	0.000
Dataset5	0.047	0.648	0.000
Dataset6	1.000	0.246	0.034

4.4.1 Problem Formulation

The prediction of biochar yield is considered a non-linear regression problem, as there are several parameters that are influenced by the production of biochar. For example, the dataset used in this study was taken from Ref. [8] and consists of 33 records; each record has five operating parameters. These parameters are the pyrolysis temperature ($^{\circ}\text{C}$), holding time (min), heating rate ($^{\circ}\text{C min}^{-1}$), percentage of moisture content (%), and sample mass (g), in which the pyrolysis temperature (x_1) is constrained in the range 400°C to 600°C , heating rate (x_2) in the range 4 to $16^{\circ}\text{C min}^{-1}$, holding time (x_3) in the range 40 – 100 min, moisture content (x_4) in the range 45 to 85% , and sample mass (x_5) in the range 5 – 20 g. The objective function fit that maximizes biochar production is defined as

$$y = \max_{x_1, x_2, x_3, x_4, x_5} fit(x_1, x_2, x_3, x_4, x_5), \text{ subject to, } 400 \leq x_1 \leq 600, \\ 45\% \leq x_2 \leq 85\%, 40 \leq x_3 \leq 100, 5 \leq x_4 \leq 20, 4 \leq x_5 \leq 16. \quad (24)$$

4.4.2 Results and Discussion

To evaluate the ANFIS-GWO in predicting biochar yield, it is compared with the original ANFIS and seven optimized ANFIS with different meta-heuristic algorithms, i.e. PSO, GA, grasshopper optimization algorithm (GOA) [43], sine-cosine algorithm (SCA) [34], whale optimization algorithm (WOA) [35], and flower pollination algorithm (Flower) [52], as well as LS-SVM and regression NN [8]. These algorithms were selected because they achieved good results in previous works [2, 11, 12, 14, 15, 19]. The parameter settings of these algorithms are listed in Table 4.

The first step in this experiment is to determine the optimal number of membership functions (clusters) by using the FCM method. Therefore, the FCM is applied at different cluster number values, and the results

Table 4: Parameter Setting of the Algorithms.

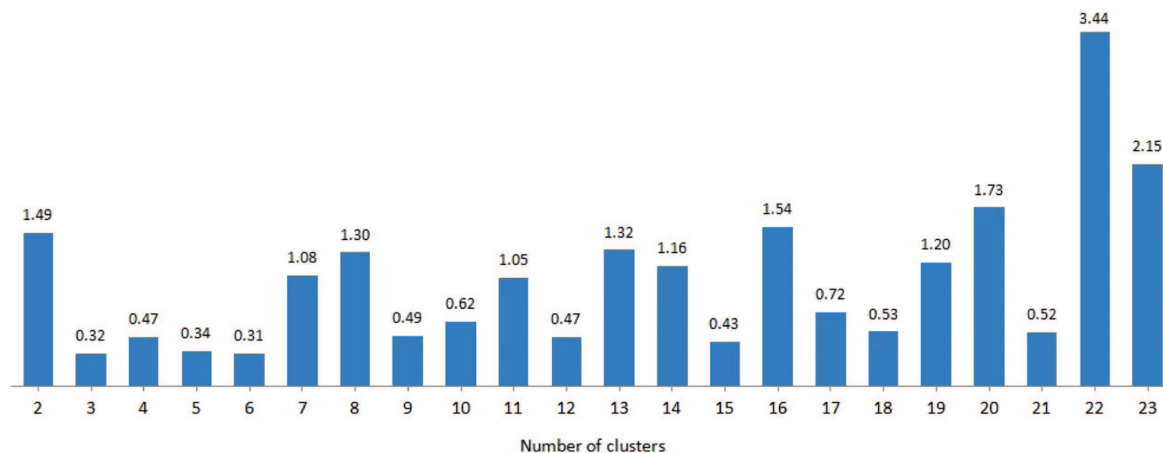
Algorithm	Parameter values
GWO	$a = [2 : 0]$
GA	$pc = 0.8, \gamma = 0.2, pm = 0.3, \mu = 0.02, \beta = 8$
PSO	$w = 1, wDamp = 0.99, C1 = 1, C2 = 2$
GOA	$c_{max} = 1, c_{min} = 0.00004$
SCA	$a = 2$
WOA	$a = [0, 2], b = 1, l = [-1, 1]$
Flower	Proximity probability = 0.8
ANFIS	ErrorGoal = 0, initial step size = 0.01

are shown in Figure 8. From this figure, it can be concluded that the optimal numbers of memberships are 3 and 6 with RMSE 0.32 and 0.31, respectively; hence, we set the number of memberships to 6.

The results of the proposed method against nine models are presented in Tables 5 and 6 and Figures 9 and 10, which are the average of 15 independent runs. From Table 5, it can be concluded that, first, when the proposed method was compared with the original ANFIS model and the optimized ANFIS models, it had the best values of RMSE and AAPRE. Therefore, its performance is >35% of the original ANFIS. Second, when the proposed method was compared with five optimized ANFIS models, regression NN, and LS-SVM, it also had the best results in terms of all measures.

Moreover, in terms of R^2 , the result of the proposed method was better than that of all other algorithms, as shown in Figures 9 and 10. These figures indicate that the results of ANFIS-GWO were nearest to the target data compared with the original ANFIS and all other models.

To test the stability of the models, the standard division measure was calculated and listed in Table 5. From this table, it can be seen that the optimized ANFIS with GWO had the best value and showed a high stability against the other models.

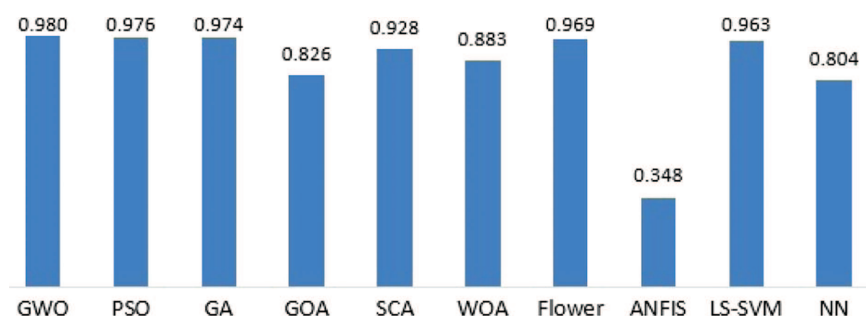
**Figure 8:** Results to Determine the Optimal Number of Clusters of FCM.**Table 5:** Performance of the Proposed Method against Other Algorithms.

	GWO	PSO	GA	GOA	SCA	WOA	Flower	ANFIS	LS-SVM	NN
RMSE	0.259	0.263	0.263	0.388	0.294	0.311	0.307	0.720	0.365	0.835
AAPRE	4.525	4.688	6.106	7.967	5.791	5.981	5.965	4.945	4.945	9.662
R^2	0.980	0.976	0.974	0.826	0.928	0.883	0.969	0.348	0.963	0.804
SD	0.0463	0.0479	0.0478	0.0839	0.0494	0.0519	0.0767	0.1824	—	—
Time	9.857	10.118	11.373	180.55	9.859	10.685	10.110	—	—	—

Boldface indicates the best result.

Table 6: Comparison between the Sample Case of Predicted Data and Target Data.

System output (predicted data)									Target data
Improved ANFIS using									
GWO	PSO	GA	GOA	SCA	WOA	Flower	ANFIS	LS-SVM	NN
3.995	4.199	4.011	4.240	3.897	3.848	4.018	3.952	3.990	3.870
7.732	7.459	7.638	7.581	7.587	7.478	7.659	6.207	7.390	5.410
7.014	7.877	6.942	6.776	6.814	7.015	6.951	6.288	6.700	7.360
5.631	5.596	5.593	5.189	5.604	5.755	5.558	6.432	5.590	4.770
5.183	4.972	5.121	5.273	5.196	5.055	5.171	6.070	4.900	6.550
3.317	3.383	3.246	3.132	3.278	3.138	3.228	2.876	3.260	3.790
1.653	1.727	1.814	2.209	1.990	2.051	1.848	1.749	2.150	1.950
4.554	4.464	4.516	4.437	4.828	4.488	4.537	4.857	4.490	4.860
3.670	3.897	3.658	3.645	3.301	3.435	3.525	3.624	3.700	3.690
7.974	8.677	7.786	7.240	7.535	7.550	7.733	6.627	7.400	7.530

**Figure 9:** Accuracy of All Algorithms According to R^2 Measure.

From the above results, in general, it can be noticed that when meta-heuristic techniques were used to learn the parameters of ANFIS, the performance of the ANFIS model improved in terms of RMSE, R^2 , and SD.

In terms of computational time, the proposed model had the lowest run time than other models due to the ability of GWO to explore and exploit the population in less time than the other algorithms.

For a deeper analysis of the performance of the proposed method, the convergence curves were considered and illustrated in Figure 11 to show the convergence behavior of the algorithms. From this figure, it can be seen that the GWO came in fourth place after GA, PSO, and WOA. Nevertheless, GWO reached the best fitness value compared with all other algorithms.

The sample cases of the predicted data are arranged in Table 6, which shows the effectiveness of the proposed method against that of the other models.

These results imply that the proposed ANFIS-GWO model improves the performance and efficiency of ANFIS by utilizing the strengths of the GWO algorithm, which led to good accuracy and stability compared with the other models.

4.5 Experiment Series 3: Estimating the Values of Input Parameters of Pyrolysis

In this experiment, the performance of the proposed ANFIS-GWO method is evaluated by using it to estimate the values of input parameters of pyrolysis that maximize the biochar production. The problem formulation in this experiment is the same as in the previous one except for the input parameters, which here are considered unknown; only the biochar yield value is given in Eq. (24). Solving the biochar production approach is a challenge; the GWO algorithm is used to solve the approach in Eq. (24). To evaluate the performance of the proposed method, the average of the RMSE for 15 runs is computed. The estimated values of the input parameters of pyrolysis are illustrated in Table 7. According to the results in Table 7, it can be concluded that the

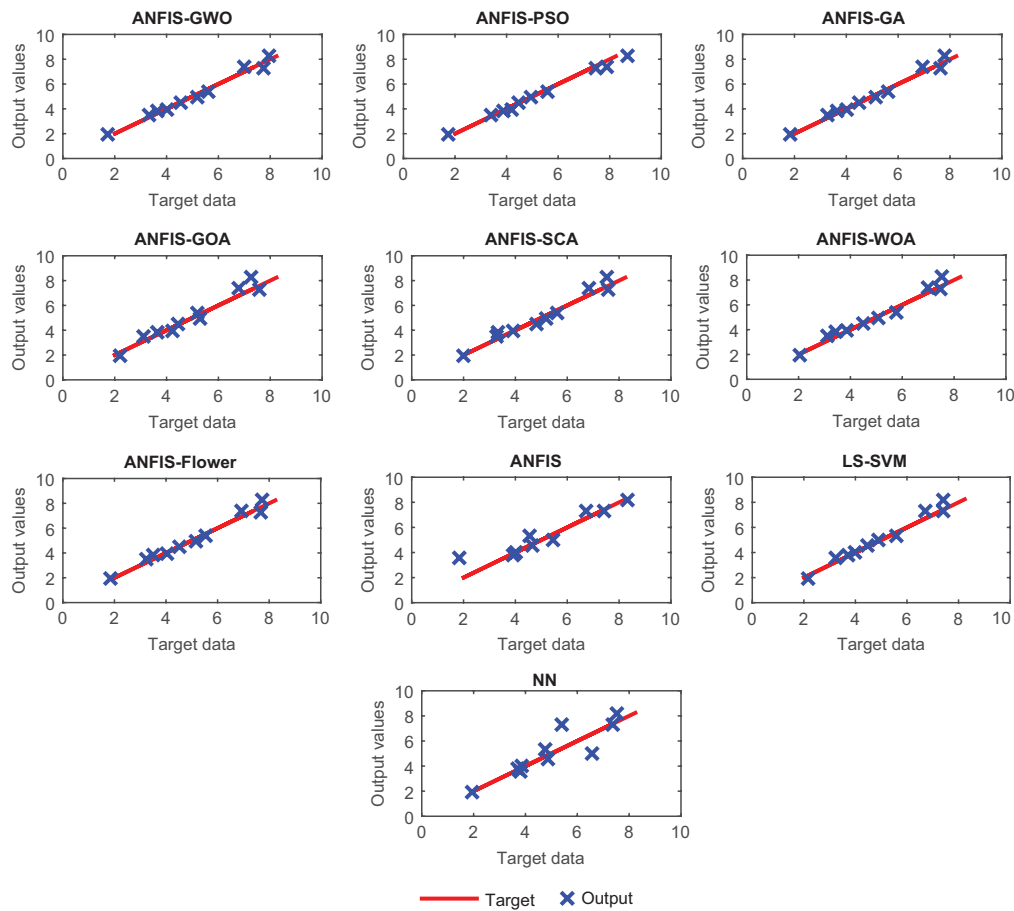


Figure 10: Predicted Data of All Algorithms against the Target (Real) Data.

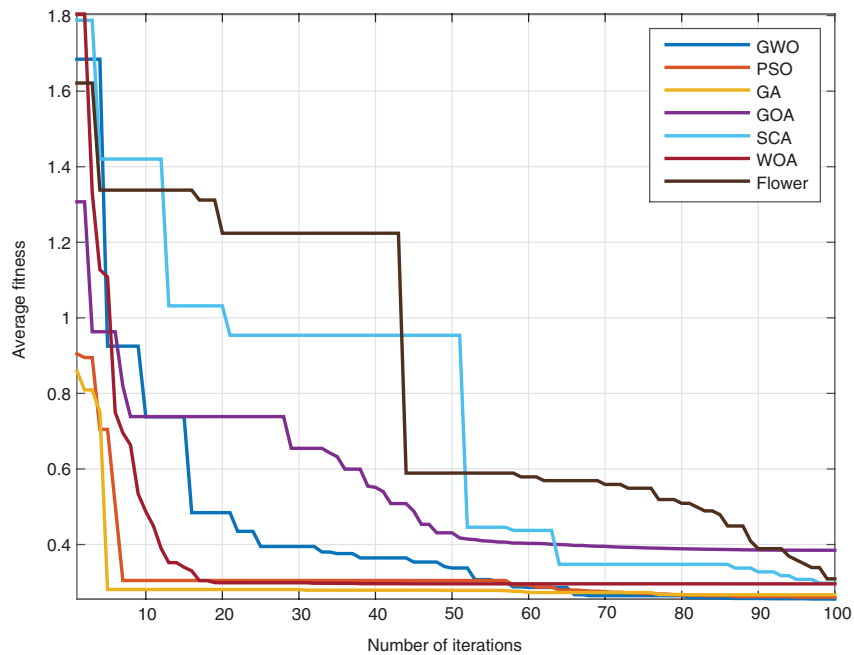


Figure 11: Convergence Curve of All Optimization Algorithms.

Table 7: Estimating the Values of Input Parameters.

	Pyrolysis temperature (°C)	Heating rate (°C min ⁻¹)	Holding time (min)	Moisture content (%)	Sample mass (g)	Biochar mass (g)	RMSE
Target	500	12	40	55	20	7.28	0.34
Estimated	500.46	11.89	39.63	55.27	19.60		
Target	500	12	60	55	10	3.9	0.49
Estimated	500.91	12.53	59.78	55.13	9.88		
Target	500	12	60	65	10	4.24	0.30
Estimated	499.45	11.78	59.89	65.31	9.97		
Target	450	8	40	85	15	5.35	0.40
Estimated	449.61	8.45	40.09	84.38	14.78		
Target	600	12	60	45	10	3.77	0.53
Estimated	600.0	12.33	60.79	44.19	10.07		

proposed ANFIS-GWO model has high accuracy to estimate the input parameters that can reach the required biochar production.

5 Conclusions and Future Works

In this study, the improvement of the ANFIS based on the GWO algorithm has been proposed to enhance the prediction performance of the biochar yield, where this process is considered an optimization problem. Therefore, several machine learning methods have been proposed to solve this problem, such as NN and LS-SVM. All of these previous methods have their shortcomings that will affect the performance of the prediction. Thus, the proposed ANFIS-GWO model avoids these limitations and was used to improve the performance of the prediction of the biochar yield from manure pyrolysis. The experimental results showed that the ANFIS-GWO model for the UCI dataset is better than other algorithms in terms of the performance measure. Meanwhile, for the prediction of biochar yield, the proposed ANFIS-GWO had the best values of RMSE and AAPRE, which were 0.259 and 4.525, respectively. Therefore, its accuracy is >35% of the standard ANFIS, and its results are also better than those of ANFIS-PSO, ANFIS-GA, LS-SVM, and regression NN. Moreover, the values of input parameters that maximize the biochar production were estimated using the GWO algorithm and showed good results.

Based on the promising results of the proposed model, in future works, it can be applied to different other applications such as quantitative structure-activity relationship, solar radiation, and biogas production.

Bibliography

- [1] K. Ahmed, A. A. Ewees, M. A. El Aziz, A. E. Hassanien, T. Gaber, P. W. Tsai and J. S. Pan, A hybrid krill-ANFIS model for wind speed forecasting, in: *International Conference on Advanced Intelligent Systems and Informatics*, pp. 365–372, Springer, Cham, 2016.
- [2] K. Ahmed, A. A. Ewees and A. E. Hassanien, Prediction and management system for forest fires based on hybrid flower pollination optimization algorithm and adaptive neuro-fuzzy inference system, in: *Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)*, IEEE, pp. 305–310, 2017.
- [3] S. An, H. Shi, Q. Hu, X. Li and J. Dang, Fuzzy rough regression with application to wind speed prediction, *Inform. Sci.* **282** (2014), 388–400.
- [4] C. J. Atkinson, J. D. Fitzgerald and N. A. Hipps, Potential mechanisms for achieving agricultural benefits from biochar application to temperate soils: a review, *Plant Soil* **337** (2010), 1–18.
- [5] H. Basser, H. Karami, S. Shamshirband, S. Akib, M. Amirmojahedi, R. Ahmad, A. Jahangirzadeh and H. Javidnia, Hybrid ANFIS-PSO approach for predicting optimum parameters of a protective spur dike, *Appl. Soft Comput.* **30** (2015), 642–649.
- [6] P. Berkhin, *A Survey of Clustering Data Mining Techniques*, Springer, Berlin, pp. 25–71, 2006.
- [7] J. C. Bezdek, R. Ehrlich and W. Full, FCM: the fuzzy c-means clustering algorithm, *Comput. Geosci.* **10** (1984), 191–203.
- [8] H. Cao, Y. Xin and Q. Yuan, Prediction of biochar yield from cattle manure pyrolysis via least squares support vector machine intelligent approach, *Bioresour. Technol.* **202** (2016), 158–164.

- [9] F. Chekired, A. Mellit, S. A. Kalogirou and C. Larbes, Intelligent maximum power point trackers for photovoltaic applications using FPGA chip: a comparative study, *Solar Energy* **101** (2014), 83–99.
- [10] A. Dukpa, I. Duggal, B. Venkatesh and L. Y. Chang, Optimal participation and risk mitigation of wind generators in an electricity market, *IET Renew. Power Gen.* **4** (2010), 165–175.
- [11] M. A. El Aziz, A. A. Ewees and A. E. Hassanien, Whale Optimization Algorithm and Moth-Flame Optimization for multilevel thresholding image segmentation, *Expert Syst. Appl.* **83** (2017), 242–256.
- [12] M. E. A. Elaziz, A. A. Ewees, D. Oliva, P. Duan and S. Xiong, A hybrid method of sine cosine algorithm and differential evolution for feature selection, in: *International Conference on Neural Information Processing*, pp. 145–155, Springer, Cham, 2017.
- [13] M. A. El Aziz, A. M. Hemdan, A. A. Ewees, M. Elhoseny, A. Shehab, A. E. Hassanien and S. Xiong, Prediction of biochar yield using adaptive neuro-fuzzy inference system with particle swarm optimization, in: *2017 IEEE PES PowerAfrica*, pp. 115–120, IEEE, 2017.
- [14] M. A. El Aziz, A. A. Ewees and A. E. Hassanien, Multi-objective whale optimization algorithm for content-based image retrieval, *Multimed. Tools Appl.* **77** (2018), 26135–26172.
- [15] M. A. El Aziz, A. A. Ewees, A. E. Hassanien, M. Mudhsh and S. Xiong, Multi-objective whale optimization algorithm for multilevel thresholding segmentation, in: *Advances in Soft Computing and Machine Learning in Image Processing*, pp. 23–39, Springer, Cham, 2018.
- [16] A. M. El-Zonkoly, A. A. Khalil and N. M. Ahmied, Optimal tuning of lead-lag and fuzzy logic power system stabilizers using particle swarm optimization, *Expert Syst. Appl.* **36** (2009), 2097–2106.
- [17] E. Emary, H. M. Zawbaa and C. Grosan, Experienced gray wolf optimization through reinforcement learning and neural networks, *IEEE Trans. Neural Netw. Learn. Syst.* **29** (2018), 681–694.
- [18] A. A. Ewees, M. A. El Aziz and M. Elhoseny, Social-spider optimization algorithm for improving ANFIS to predict biochar yield, in: *8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE, pp. 1–6, 2017.
- [19] A. A. Ewees, M. A. El Aziz and A. E. Hassanien, Chaotic multi-verse optimizer-based feature selection, *Neural Comput. Appl.* in press. (2017), 1–16. DOI: <https://doi.org/10.1007/s00521-017-3131-4>.
- [20] H. Faris, I. Aljarah, M. A. Al-Betar and S. Mirjalili, Grey wolf optimizer: a review of recent variants and applications, *Neural Comput. Appl.* **30** (2018), 413–435.
- [21] N. P. Gurwick, L. A. Moore, C. Kelly and P. Elias, A systematic review of biochar research, with a focus on its stability in situ and its promise as a climate mitigation strategy, *PLoS One* **8** (2013), e75932.
- [22] International Biochar Initiative, Standardized product definition and product testing guidelines for biochar that is used in soil, in: *IBI Biochar Standards*, 2012.
- [23] J. S. R. Jang, ANFIS: adaptive-network-based fuzzy inference system, *IEEE Trans. Syst. Man Cybernet.* **23** (1993), 665–685.
- [24] Y. A. Katsigiannis, P. S. Georgilakis and E. S. Karapidakis, Multiobjective genetic algorithm solution to the optimum economic and environmental performance problem of small autonomous hybrid power systems with renewables, *IET Renew. Power Gen.* **4** (2010), 404–419.
- [25] A. Kumar, S. Pant and M. Ram, System reliability optimization using gray wolf optimizer algorithm, *Qual. Reliab. Eng. Int.* **33** (2016), 1327–1335.
- [26] J. Lehmann, A handful of carbon, *Nature* **447** (2007), 143–144.
- [27] D. F. Lekkas, C. E. Imrie and M. J. Lees, Improved non-linear transfer function and neural network methods of flow routing for real-time forecasting, *J. Hydroinform.* **3** (2001), 153–164.
- [28] M. Lichman, *UCI Machine Learning Repository*, 2013.
- [29] G. H. Lin, X. Y. Lou and B. T. Cui, Parameter design of LS-SVM based on QPSO and its application to node localization, *Appl. Mech. Mater.* **490** (2014), 542–545.
- [30] E. G. Mansoori, M. J. Zolghadri and S. D. Katebi, SGERD: a steady-state genetic algorithm for extracting fuzzy classification rules from data, *IEEE Trans. Fuzzy Syst.* **16** (2008), 1061–1071.
- [31] J. J. Manyá, Pyrolysis for biochar purposes: a review to establish current knowledge gaps and research needs, *Environ. Sci. Technol.* **46** (2012), 7939–7954.
- [32] A. P. McCabe, G. A. Aggidis and M. B. Widden, Optimizing the shape of a surge-and-pitch wave energy collector using a genetic algorithm, *Renew. Energy* **35** (2010), 2767–2775.
- [33] A. Mellit, S. A. Kalogirou and M. Drif, Application of neural networks and genetic algorithms for sizing of photovoltaic systems, *Renew. Energy* **35** (2010), 2881–2893.
- [34] S. Mirjalili, SCA: a sine cosine algorithm for solving optimization problems, *Knowl.-Based Syst.* **96** (2016), 120–133.
- [35] S. Mirjalili and A. Lewis, The whale optimization algorithm, *Adv. Eng. Softw.* **95** (2016), 51–67.
- [36] S. Mirjalili, S. M. Mirjalili and A. Lewis, Grey wolf optimizer, *Adv. Eng. Softw.* **69** (2014), 46–61.
- [37] D. Oliva, A. A. Ewees, M. A. E. Aziz, A. E. Hassanien and M. Peréz-Cisneros, A chaotic improved artificial bee colony for parameter estimation of photovoltaic cells, *Energies* **10** (2017), 865.
- [38] A. Parsian, M. Ramezani and N. Ghadimi, A hybrid neural network-gray wolf optimization algorithm for melanoma detection, *Biomed. Res.* **28** (2017), 3408–3411.

- [39] M. Puig-Arnavat, J. C. Bruno and A. Coronas, Review and analysis of biomass gasification approaches, *Renew. Sustain. Energy Rev.* **14** (2010), 2841–2851.
- [40] K. Qian, A. Kumar, H. Zhang, D. Bellmer and R. Huhnke, Recent advances in utilization of biochar, *Renew. Sustain. Energy Rev.* **42** (2015), 1055–1064.
- [41] Z. Ramedani, M. Omid, A. Keyhani, B. Khoshnevisan and H. Saboohi, A comparative study between fuzzy linear regression and support vector regression for global solar radiation prediction in Iran, *Solar Energy* **109** (2014), 135–143.
- [42] I. Robandi, Photovoltaic parameter estimation using Grey Wolf Optimization, in: *3rd International Conference on Control, Automation and Robotics (ICCAR)*, IEEE, pp. 593–597, 2017.
- [43] S. Saremi, S. Mirjalili and A. Lewis, Grasshopper optimisation algorithm: theory and application, *Adv. Eng. Softw.* **105** (2017), 30–47.
- [44] M. Shahlaei, A. Madadkar-Sobhani, L. Saghaie and A. Fassihi, Application of an expert system based on Genetic Algorithm-Adaptive Neuro-Fuzzy Inference System (GA-ANFIS) in QSAR of cathepsin K inhibitors, *Expert Syst. Appl.* **39** (2012), 6182–6191.
- [45] K. D. Sharma, A. Chatterjee and A. Rakshit, A hybrid approach for design of stable adaptive fuzzy controllers employing Lyapunov theory and particle swarm optimization, *IEEE Trans. Fuzzy Syst.* **17** (2009), 329–342.
- [46] B. Sivakumar, A. W. Jayawardena and T. M. K. G. Fernando, River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches, *J. Hydrol.* **265** (2002), 225–245.
- [47] I. Stavi and R. Lal, Agroforestry and biochar to offset climate change: a review, *Agron. Sustain. Dev.* **33** (2013), 81–96.
- [48] L. Suganthi, S. Iniyar and A. A. Samuel, Applications of fuzzy logic in renewable energy systems – a review, *Renew. Sustain. Energy Rev.* **48** (2015), 585–607.
- [49] K. Tselepidou and K. L. Katsifarakis, Optimization of the exploitation system of a low enthalpy geothermal aquifer with zones of different transmissivities and temperatures, *Renew. Energy* **35** (2010), 1408–1413.
- [50] J. Twidell and T. Weir, *Renewable energy resources*, Routledge, Abingdon, Oxon, 2015.
- [51] P. Weiland, Biogas production: current state and perspectives, *Appl. Microbiol. Biotechnol.* **85** (2010), 849–860.
- [52] X. S. Yang, Flower pollination algorithm for global optimization, in: *International Conference on Unconventional Computing and Natural Computation*, pp. 240–249, Springer, Berlin, 2012.