Artificial Neural Network (ANN) trained by a Novel Arithmetic Optimization Algorithm (AOA) for Short Term Forecasting of Wind Power

Muhammad Hamza Zafar[1], Noman Mujeeb Khan[1], Syed Kumayl Raza Moosavi[2], Majad Mansoor[3], Adeel Feroz Mirza[3] and Naureen Akhtar[4] (✉)

[1] Capital University of Science and Technology, Islamabad, 44000, Pakistan
[2] National University of Sciences and Technology, Islamabad, 44000, Pakistan
[3] University of Science and Technology of China, Heifi, 230027 China
[4] Dept. of Engineering Sciences, University of Agder (UiA), Jon Lilletuns vei 9, 4879, Grimstad, Norway
Naureen.Akhtar@uia.no

Abstract. Stochastic nature of wind power with high amount of non-linearity makes it very difficult to predict wind power production in real time which has a high impact in renewable energy industry. The uncertainty of wind power makes it challenging to integrate it with the power grid. As a solution, an early short term forecasting of the wind flow significantly improves the wind power generation. For this purpose, a novel arithmetic optimization algorithm is used to train an artificial neural network for short term wind power prediction. Effective exploration and exploitation behavior due to embedded arithmetic operators for updating the position effectively trains the neural network. To validate the performance of the proposed technique well-known methods are compared using case studies. The proposed method has shown better prediction performance as compared to existing techniques. AOANN achieves up to 94.87% and 97.18% less training error and up to 96.42% and 83.64% less testing error in winter and summer seasons respectively.

Keywords: Bio-inspired Neural Network, Intelligent Control System, Arithmetic Optimization Algorithm, Wind Power, Regression.

1 Introduction

The euphoria of living in a society that is technologically advanced is compelling However, this advancement has exponentially risen the demand of energy as well. Heavy reliance of modern world on power has sky-rocketed the demand for energy globally. Conventional energy resources such as oil, gas and coal are not only fatal for the living things but there are also hazardous for the environment. Now more than ever the need for alternate sources of energy has grown rapidly. Sources of energies such as wind, solar, and biogas etcetera on the other hand being clean and in-exhaustible provide alternate. These sources neither produce green gases nor can they be depleted.

Amongst the mentioned alternate energy sources, wind energy is one of the fastest growing technology. A single wind turbine can generate up-to several $MW's$ power. Wind power however is intermittent in nature, that is, it is dependent upon weather conditions. This dependency poses certain challenges ranging from economy to dispatching of wind power from power grid etcetera. Moreover, the intermittency of wind power has adverse effects on the stable operation of wind turbines. The non-stable operation of wind turbine hinders the large scale integration of wind power due to voltage and frequency fluctuations. Therefore a balance is required between power generation and transmission. For this purpose wind prediction plays a vital role in smooth and cost effective operation of wind turbines.

In order to extract the maximum power out of the wind systems, a lot of research has been carried out in the recent years. Typically two classification techniques can be found in the literature for the prediction of wind power. The first method being physical is excellent in terms of long term power prediction. However they suffer from low precision. The other method known as statistical based methods have the ability to co-relate between the wind power and corresponding input variables such as meteorological data.

Physical models, also known as deterministic models are dependent upon the numerical weather conditions, that is these models use equations of atmospheric motions in order to calculate what meteorological measurements will be in the future. These models are comprised of several features namely climate, terrain, or atmospheric condition etcetera and wind power is predicted via one of these features.

Physical models predict the power in two stages that is first, the wind speed is predicted and secondly, it is then converted into electrical power. Physical models have shown promising long term prediction results, however they suffer from certain drawbacks as well such as being costly, and difficult to design which renders them unsuitable for predicting wind power.

To overcome the short comings of physical model, statistical models have been gaining popularity are being widely researched. Statistical models make use of training datasets in order to predict the wind power. One technique such as [1] utilizes the least squares support vector machine to predict wind power. Further in order to improve the prediction results, an optimization technique known as gravitational search procedure is implemented. In some cases, datasets contain missing data. To overcome this situation, Gaussian process regression and multiple imputations [2] is introduced for wind power prediction. In this process, expectation-maximization procedure is used for estimating mixture components of the data distribution for handling the missing data. Another approach uses hybrid model to producing better wind prediction results [3]. The trend of wind is captured using the wind power curve. This trend is adjusted using data-driven schemes. One obvious problem with this method is the increased complexity and longer prediction time.

With the popularity and success in the domain of machine learning, machine learning algorithms have gained significant recognition primarily due to their performance in the recent years. These algorithms are being used by the researchers for predicting wind power, such as K-nearest neighbors (KNN) [4] which uses multiple feature meteorological input data. One robust algorithm genetic programming based ensemble of neural networks [16] has the tendency to predict wind power on short term basis. Support vector machine [5] finds a co-relation between wind speed and power by altering the invalid initial measurements. However, SVM doesn't have the ability to predict the wind power in long terms. [6] A hybrid of SVM uses a combination of wavelet transform and SVM. These hybrid algorithms have the tendency to produce exceptional results. Another modified version of SVM that uses a combination of autoregressive moving average (ARMA) model, the support vector machine (SVM) prediction, and particle swarm optimization (called a hybrid PSO-SVM-ARMA) [7] is employed to improve the prediction results significantly. [8] K-Means-long short-term memory (K-Means-LSTM) network model is capable of handling the time dependencies on a time series data which makes it superior than back-propagation neural networks, and support vector machine models.

To overcome the above mention short comings and effectively predict the wind power in real time, a novel arithmetic optimization algorithm (AOA) based artificial neural network (ANN) [9] is proposed. In this AOA-NN the algorithm is used to train the neural network model on the training dataset. The effective exploration and exploitation capability of AOA with less number of tuning parameters and less random numbers makes it efficient for the minimization of cost function during the training of neural network.

2 Meta Heuristic Algorithms

Recent years have brought forth a new set of algorithms known as meta-heuristic algorithms that are possessed with excellent capabilities in terms of locating the global minimum. Literature shows a wide variety of such algorithms, therefore, they can be split into two categories:

2.1 Single Based Agent:

Single Agent Based (SAB) algorithms make use of single agent or a single candidate in order to locate the minimum solution. Algorithms such as and not limited to, for example, Simulated Annealing (SA) [10], Greedy randomized adaptive search (GRAP) [11] which are also known as nature inspired techniques are utilized for randomizing the transfer function. This single agent is improved within the search space until an optimum solution hasn't been achieved.

2.2 Multi-Agent Based:

One of the most obvious problems associated with the single search agent is that it is limited in its searching capabilities since a single agent can only search limited number of instances for finding the desired results. This drawback of single agent technique renders it in-effective. In order to mitigate this problem, multi-agents are utilized.

Multi-agents have a tendency to find the global minimum via learning from each other's position whilst utilizing a complex network of relations. Group Teaching Optimization Algorithm (GTOA) [12], swarm intelligence algorithms such as Particle Swarm Optimization (PSO) [13], the Grey Wolf Optimization (GWO) [14], the Grasshopper Algorithm (GHO) [15], the Firefly Algorithm [16], Barnacle optimization algorithm (BMO) [17] are some of the algorithms that make use of multi-agents.

The use of these multi-agent based algorithms can also be seen in the field of deep learning. Updation of weights and biases of an ANN is accomplished through the use of such heuristic algorithms. Several algorithms have already been employed for the purpose of optimizing a neural network but since it can never be certified that a single algorithm is best for every problem, more and more algorithms are introduced every day. In this perspective, a novel, multi-agent based, arithmetic optimization algorithm [9] has been chosen as the focus of this research for the purposes of training an ANN, namely the Arithmetic Optimization Algorithm based Neural Network (AOANN).

3 Arithmetic Optimization Algorithm (AOA)

In this paper the arithmetic optimization algorithm is used to train the feed forwards neural network. This algorithm is also a population based optimization algorithm which uses arithmetic operators to update the position without calculating their derivatives. Arithmetic is a basic part of number theory but also the important part of modern mathematics. Traditional measures are used to study the numbers are arithmetic operators (that is, addition, subtraction, multiplication, and division). The inspiration for AOA comes from the use of these arithmetic operators in solving the arithmetic problems. The proposed algorithm is explained below.

3.1 Initialization:

The initialized set of candidate solutions (D), which are generated randomly as presented in (1).

$$D = \begin{bmatrix} d_{1,1} & \cdots & d_{1,j} \\ d_{2,1} & & \vdots \\ \cdots & \ddots & \vdots \\ d_{N-1,1} & & d_{N-1,j} \\ d_{N,1} & \cdots & d_{N,j} \end{bmatrix} \tag{1}$$

At the start of AOA, first needs to select the search phase (that is, exploitation or exploration) which can be done by math optimizer coefficient (MOA) function calculation using (2).

$$MOA = Min + it \times (\frac{Max - Min}{Max - it}) \tag{2}$$

where Max , Min are the maximum and minimum value of MOA. it is the current iteration while max_it represents the maximum number of iterations.

3.2 Exploration Phase:

In the arithmetic operators, high distributed values can be achieved by using multiplication M or division D operator in mathematical calculations. This leads to exploration search mechanism. Due to high dispersion created by D and M operator, it is difficult to approach target but in exploitation phase S and A operator will reach the target.

The search phases in AOA are controlled by MOA. In the exploration phase the updating of position occurs using (3) with D and M operator. If $r_1 > MOA$ then exploration phase occurs as depicted in Fig. 1.

$$d_{i,j}(it+1) = \begin{cases} best(d_j) \div (MOP + \varepsilon) \times \left((UB - LB) \times \mu + LB\right), & r_2 > 0.5 \\ best(d_j) \times (MOP) \times \left((UB - LB) \times \mu + LB\right), & otherwise \end{cases} \tag{3}$$

where best(d_j) is the global position, UB and LB are the upper and lower boundary search space, ε is the small value, μ is a control parameter which is used to adjust the search process. MOP is the math optimization probability whose value will be updated using (4)

$$MOP(it) = 1 - \frac{it^{1/\alpha}}{Max\_it^{1/\alpha}} \tag{4}$$

where α is the sensitive parameter and defines the accuracy of the exploitation over the iteration.

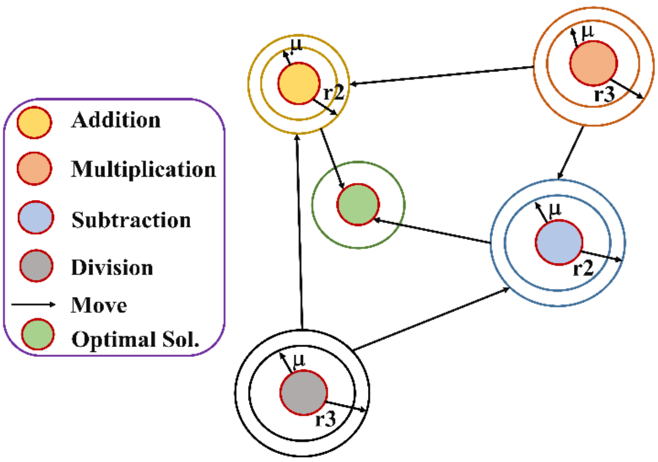

Fig. 1. Position updation of math operators in search of optimal solution.

3.3 Exploitation Phase:

In the arithmetic operators, high dense results are produced using the addition A and subtraction S operator which is the exploitation mechanism. Since S and A cause low dispersion which leads to the target value.

The condition of exploitation phase is also related to the MOA. This phase occurs when $r_1 < MOA$. The subtraction searching strategy and the addition searching strategy are used to search deeply on the dense regions which is modeled in (5).

$$d_{i,j}(it) = \begin{cases} best(d_j) - MOP \times ((UB - LB) \times \mu + LB), & r_3 > 0.5 \\ best(d_j) + MOP \times ((UB - LB) \times \mu + LB), & otherwise \end{cases} \quad (5)$$

This mechanism is modeled for deep search. The subtraction S mechanism will be implemented if $r_3 > 0.5$, where r_3 is the random number between [0,1]. If $r_3 \le MOA$ then the second operator, that is, addition A will perform the required task. The most important parameter is μ which needs to be carefully adjusted for the best stochastic process.

3.4 Pseudo Code:

The pseudo code of AOA algorithm is shown in Fig. 2. Since every solution updates its position according to the best extracted result, and also in order to focus on the exploration and exploitation, MOA increases linearly from 0.2 to 0.9.

```
initialize the AOA parameters, that is, μ, α
initialize the random population d_i = {1,2,3,⋯,N}
    while(it<max_it)
    calculate fitness of every solution
    update value of MOA and MOP using (2) and (4)
        generate random number r_1, r_2, r_3
        if r_1 > MOA
            if r_2 > 0.5
                update position using the first rule in (3)
            else
            update position using the second rule in (3)
            end if
        if r_3 > 0.5
            update position using the first rule in (5)
        else
            update position using the second rule in (5)
        end if
        end if
        it=it+1
    end while
Return the best solution d
```

Fig. 2. Pseudo Code of the propsed AOANN for short term wind forcasting updation.

4 Design of AOA-NN

In this paper a 3 layer neural network, that is, input, hidden, and output layer is proposed. The hidden layer contains 10 neurons. The proposed structure of NN is shown in Fig. 3. As shown in the AOA algorithm is used to update the weights and biases.

The input vector contains the instances of N input features and the corresponding output y_i. The AOA-NN is implemented in the MATLAB 2018a. The flowchart for updating the weights and the biases are shown in Fig. 4.

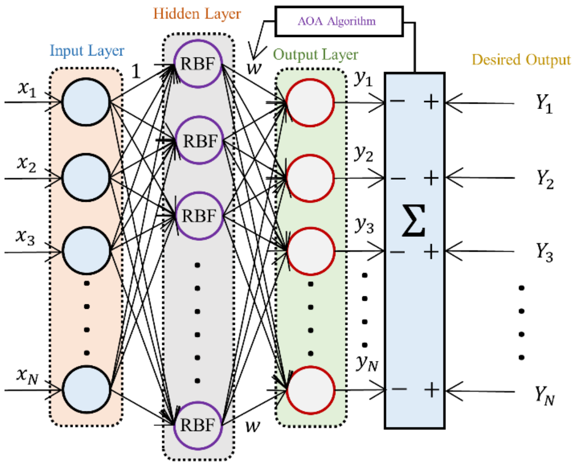

Fig. 3. Proposed Structure of AOA0-NN.

4.1 Neural Network Design:

Fig. 3 depicts the structure that is used to design the neural network. Selection for number of neurons is needed to be made or higher efficiency with less computational time. Neural network with low number of neurons and hidden layers may not be able to predict the output very efficiently. Neural network with high number of neurons and hidden layers which results in high efficiency but the computational complexity will increase drastically. Therefore, the right sets of variables are required for highest order of accuracy.

4.2 Activation Function:

In order to effectively train and test the neural network model, the right activation function for the hidden layer is required. For classification of dataset, the sigmoid function is suitable as shown in (6).

$$\hat{a_i} = \frac{1}{1 + e^{-x_i}} \tag{6}$$

Since wind power prediction is a regression problem, meaning the neural network will predict the continuous values. Therefore, the activation function used for this problem is radial basis function as shown in (7) and (8).

$$h(x) = e^{-(\frac{(x-c)^2}{r^2})} \tag{7}$$

$$y(x) = \sum_{j=1}^{N} w_j h_j(x) \tag{8}$$

where $h(x)$ is the function for the hidden layer and $y(x)$ is the predicted output.

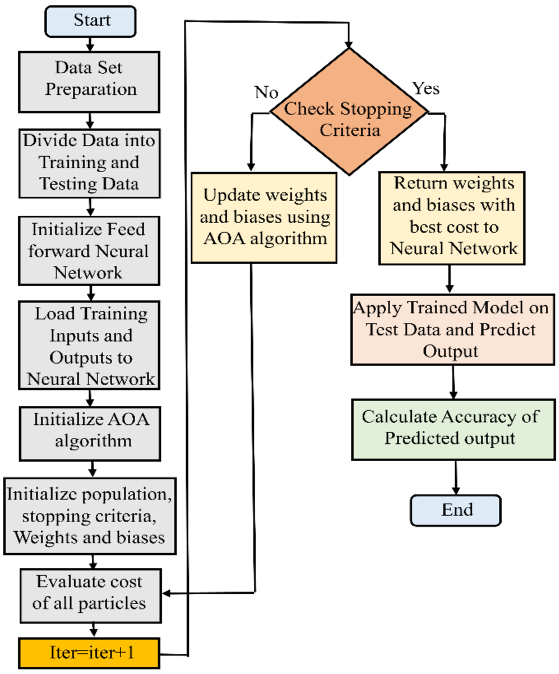

Fig. 4. AOA-NN flowchart.

4.3 Fitness Function:

In optimization problem the most important part is the definition of the cost function which is also known as a fitness function. The weights and the biases of the neural network needs to be updated in such a way that the cost function is minimized in the training process. When the cost function is minimized, the latest updated weights and the biases will be give the best relation between input and output. The fitness function (*F.F.i*) chosen for the neural network is the normalized mean square error presented in (9).

$$F.F.i = \frac{1}{N}\sum_{j=1}^{N}(Y_j - Y_j^n)^2 \quad (9)$$

where Y_j is the true output while Y_j^n is the predicted output.

5 Results and Discussions

In this section the training and testing results of model are discussed. The comparison of AOANN is made with PSONN, GWONN and BMONN. The statistical analysis, training error and testing error suggest the superior performance of the proposed technique.

5.1 Preparation of Dataset:

In wind turbine system, SCADA systems are utilized to record the wind direction, the wind speed and power generated by the wind turbine [18]. The dataset taken is the wind turbine located in Turkey through SCADA systems with 10 min intervals. The dataset is also available at [19].

5.2 Forecasting of Wind Power:

The performance of proposed AOANN is compared with POSNN [20], GWONN [21] and BMONN. Comparison is made with true wind power and predicted wind power by all four techniques. The data is dividing into training and testing by ratio of 67% and 33% respectively. First all techniques are trained on training data with optimally tuned parameters. Then proposed method and other comparative techniques are tested on testing data. Table. I shows the minimum error achieved by all techniques during training which clearly indicate that proposed techniques achieve less error as compared to other techniques. The training accuracy of comparative techniques are also presented in Fig. 5 (a). Which shows that AOANN effectively minimize cost in less epochs. Fig. 6 (a) shows that the wind power is highly non-linear and varies greatly in 48h. The power predicted by the AOANN is close to the actual curve. This shows that AOANN is highly effective for the prediction of wind power which is highly volatile. The robustness of the proposed techniques can also be verified from Fig. 6 (b) which shows the relative error generated by the proposed technique to be very less as compared to other comparative techniques. The relative error can be calculated by

$$R.E.(\%) = \left(\frac{\mathrm{Y} - \mathrm{Y}'}{\mathrm{Y}'}\right)100\% \quad (10)$$

where Y is the actual value and Y′ is predicted value.

Table 1. Comparison of NMSE of training data.

Tech.	Training Error	
	Winter	**Summer**
AOANN	0.0061	0.0040
BMONN	0.01760	0.0165
GWONN	0.0794	0.2196
PSONN	0.1191	0.1419

Table 2. Comparison of NMSE of testing data.

Tech.	Training Error	
	Winter	**Summer**
AOANN	0.0172	0.0705
BMONN	0.0855	0.1261
GWONN	0.4140	0.8016
PSONN	0.4812	2.262

The Vectoral map shows that the summer season in Turkey shows larger variations in wind torrents due to stronger winds. The performance of proposed technique is also compared for the summer season. Table 1 and table 2 shows the cost function achieved during training process for winter and summer season data. Fig. 5 (b) shows comparison of cost achieved w.r.t. epochs by competing techniques and corresponding power prediction is presented in Fig. 7 (a). The power in summer season is highly non-linear but still AOANN achieves closer prediction to the actual value. The performance of proposed techniques in summer season can also be validated by the relative error shown in Fig. 7 (b).

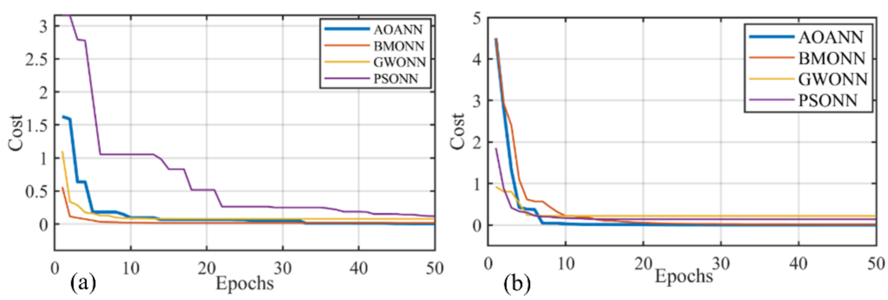

Fig. 5 (a) Cost Vs Epochs for training in Winter Season **(b)** Cost Vs Epochs for training in Summer Season.

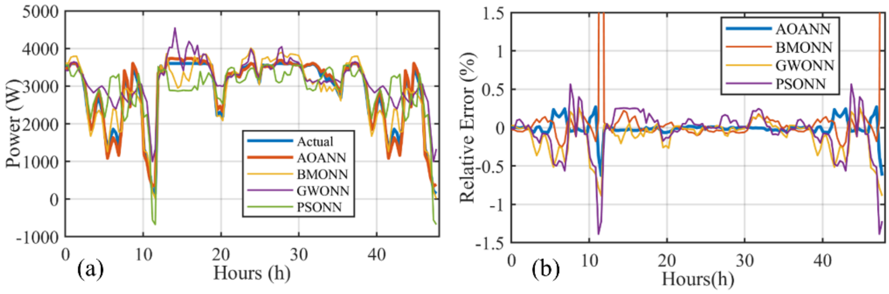

Fig. 6 (a) Comparison of Prediction of Wind Power during Winter Season **(b)** Relative Error comparison for winter season

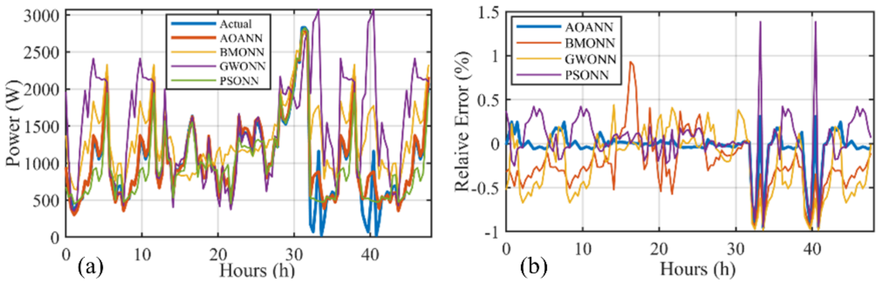

Fig. 7 (a) Comparison of prediction of wind power during summer season **(b)** Relative Error comparison for summer season

6 Conclusion

This paper presents a novel implementation of AOANN based precise forecasting of wind power promotes. The objective of the study is to improvise efficiency of the wind energy conversion systems of wind turbines and stability of grid connected operations. The random and intermittent nature of wind leads to a non-linear wind power production, leading to difficulty in achieving precise prediction and MPPT control. AOAAN based prediction model is modelled in details. Comparison is made with recently developed techniques i.e., PSONN, GWONN and BMONN. Training error, testing error, wind power prediction, relative error and efficiency parameters are utilized for comparison. To check the performance of the proposed technique in real time, model is applied on summer and winter weather conditions of Turkey. The statistical indices validate the superior performance of the proposed techniques. It is safe to conclude that AOANN is the best fit for short term wind power prediction.

References

1. X. Yuan et al., *Short-term wind power prediction based on LSSVM–GSA model.* Energy Conversion and Management 101, 393-401 (2015).

2. T. Liu, H. Wei, and K. Zhang, Wind power prediction with missing data using Gaussian process regression and multiple imputation. Appl. Soft Comput. 71, 905-916 (2018).
3. J. Yan and T. Ouyang, Advanced wind power prediction based on data-driven error correction. Energy Convers. Manage. 180, 302-311 (2019).
4. M. Yesilbudak, S. Sagiroglu, and I. Colak, A novel implementation of kNN classifier based on multi-tupled meteorological input data for wind power prediction. Energy Convers. Manage. 135, 434 444 (2017).
5. A. Zameer, J. Arshad, A. Khan, and M. A. Z. Raja, Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks. Energy Convers. Manage. 134, 361-372 (2017).
6. Y. Liu, J. Shi, Y. Yang, and W.-J. Lee, Short-term wind-power prediction based on wavelet transform support vector machine and statistic-characteristics analysis. IEEE Trans. Ind. Appl. 48(4), 1136-1141(2012).
7. Y. Wang, D. Wang, and Y. Tang, Clustered hybrid wind power prediction model based on ARMA, PSO-SVM, and clustering methods. IEEE Access 8, 17071-17079, (2020).
8. B. Zhou, X. Ma, Y. Luo, and D. Yang, Wind power prediction based on LSTM networks and nonparametric kernel density estimation. IEEE Access 7, 165279-165292, (2019).
9. L. Abualigah et al., *The arithmetic optimization algorithm.* Computer methods in applied mechanics and engineering 376, 113609 (2021).
10. A.M. Fathollahi-Fard et al., A green home health care supply chain: New modified simulated annealing algorithms. Journal of Cleaner Production 240, 118200 (2019).
11. A. Exposito-Marquez et al., Greedy randomized adaptive search procedure to design waste collection routes in La Palma. Computers & Industrial Engineering 137, 106047 (2019).
12. M.H. Zafar et al., Group Teaching Optimization Algorithm Based MPPT Control of PV Systems under Partial Shading and Complex Partial Shading. Electronics 9(11), 1962 (2020).
13. D. Wang, D. Tan, and L. Liu, Particle swarm optimization algorithm: an overview. Soft Computing. 22(2), 387-408 (2018).
14. S. Mirjalili, S.M. Mirjalili, and A. Lewis, Grey wolf optimizer. Advances in engineering software 69, 46-61 (2014).
15. S.Z. Mirjalili, et al., Grasshopper optimization algorithm for multi-objective optimization problems. Applied Intelligence 48(4), 805-820 (2018).
16. A.H. Gandomi, et al., Firefly algorithm with chaos. Communications in Nonlinear Science and Numerical Simulation 18(1), 89-98 (2013).
17. M.H. Sulaiman et al., *Barnacles mating optimizer: a new bio-inspired algorithm for solving engineering optimization problems.* Engineering Applications of Artificial Intelligence 87, 103330 (2020).
18. S. Şahin, and M. Türkeş, *Assessing wind energy potential of Turkey via vectoral map of prevailing wind and mean wind of Turkey.* Theoretical and Applied Climatology 141, 1351-1366 (2020).
19. https://www.kaggle.com/berkerisen/wind-turbine-scada-dataset
20. K. Khan, and A. Sahai, *A comparison of BA, GA, PSO, BP and LM for training feed forward neural networks in e-learning context.* International Journal of Intelligent Systems and Applications 4(7), 23 (2012).
21. L. Xu et al., *GWO-BP neural network based OP performance prediction for mobile multiuser communication networks.* IEEE Access 7, 152690-152700 (2019).