A Hybrid Forecasting Model Based on Modified Bat Algorithm and ELM: A Case Study for Wind Speed Forecasting

Yujia Zhang, Long Chen
School of Automation, Northwestern Polytechnical University, Xi’an, 710072, China
Email: zyjagllfb@163.com

Abstract. Accurate wind speed forecasting is essential to the dispatch and management of wind power systems for the improvement in operation reliability of wind power plants. However, the wind velocity is the most volatile and random nonlinear series which has difficulty to obtain satisfactory prediction values. Currently to achieve higher forecasting accuracy, some numerical optimization algorithms have been employed in computational methods to remedy the deficiencies. In this paper, we propose a hybrid forecasting model based on a modified bat algorithm and extreme learning machine (ELM). The process consists of two layers: a bat algorithm (BA) is improved by conjugate gradient method to optimize the parameter of ELM, which proves to converge faster than using bat algorithm purely. The second layer is the training of ELM network to obtain the final forecasting results. To verify the effectiveness of proposed novel model, we collect data from wind power stations in Penglai, China, with the comparison indicating that MBA-ELM algorithm gains more accurate forecasting results than purely applying ELM for forecasting.

1. Introduction
Along with the energy and environmental problem is outstanding day by day, wind power, which developed as a clean and pollution-free energy source, has gotten more attention from governments, academic institutions and researchers [1]. Generally, wind speed forecasting is a vital part in the operation of the wind power system considering the running characteristics of the wind turbine vary with the random fluctuation of the wind speed [2]. Accurate forecasting model can reduce the impact of large scale wind power integration on the security and stability of power systems and enhance the control of wind turbines. However, due to the wind velocity is the most unstable and volatile weather parameter which can be seen as the random nonlinear series that hard to predict [3].

Given its importance and difficulties, in recent years many scholars have put forward multiply methods for wind speed forecasting, which can be categorized as four types mainly: physical models, artificial intelligence models, statistical models and hybrid models [4]. Physical models usually describe the temperature, topography and the like information of the wind farm based on physical parameters, with NWP technology has been widely employed [5]. Statistical models usually involve the application of statistical equations, with one of the most employed models is the ARIMA model [6]. By regarding the predicted wind velocity as a random sequence, certain purposed mathematical models can approximate the sequence to predict future values from the historical data [7]. In artificial intelligence models, some artificial neural networks such as BP neural networks, RBF neural networks and ALE neural networks [8] are widely developed. But its initial weights are set randomly which may lead to a slow convergence of optimal value. Thus, other intelligent optimization algorithms should be applied to find the optimal initial weights of ANNs for the numerical optimization of initial parameters.
To conquer the limitation, many scholars proposed hybrid forecasting models with data-processing algorithms employed in ANNs. Among these algorithms, bat algorithm is a novel optimization algorithm proposed by Yang [9] which is widely adopted to search for the optimal weight coefficients for optimization of initial wind speed forecasting model [10]. Some literatures investigated into further achieve high convergence rates, and extreme learning machine (ELM) is a good choice [11]. Here in our paper, a hybrid model, which contains ELM and an improved BA, is proposed for multi-step wind speed forecasting. In the improved BA, which is on the basis of conjugate gradient method to improve convergence performance over time and prevent individual bats from entrapment in local optima, is introduced to optimize the initial weights and thresholds of ELM. The aim of this study is to investigate and enhance the forecasting performance of hybrid model based on intelligent optimization algorithm and artificial neural networks for multi-step accurate wind speed forecasting. To investigate the forecasting abilities of the proposed model, the wind speed data collected from four different wind power stations in Penglai, China, were used as a case study.

2. Modified Bat Algorithm

Bat algorithm (BA) is proposed by the inspiration from the echolocation behaviour of natural bats in finding their foods, BA is a novel swarm intelligence optimization algorithm that is widely applied in forecasting models. However, conventional BA usually shows slow convergence when it is applied to large-scale and complex problems [12]. To tackle its weakness, we proposed a new modified BA (based on conjugate gradient method) in local search procedure to speed up the convergence of local search. The core thought is to combine conjugation and steepest descent method, using the gradient of known points to construct conjugate directions and search along these directions.

Here in constructed algorithm, we denote yt to be the positions of bats in BA, with r represents the current iteration number. Note that compare to conventional BA, an inner circulation loop is added which applies conjugate gradient method. In this circulation, the current generation will be the initial value, with the gradient denoted as \( -\nabla f(y^i) \) and \( \lambda'_i \) is step length. Iterative loop presented as follows:

\[
y_{i+1}^t = y_i^t + \lambda_i^t \delta_i^t (i = 0, 1, \ldots, N - 1)
\]  

From the above formulation, \( \delta_i^t = -\nabla f(y_i^t) \). Parameter N represents the total number of iterations. By running the iteration loop, the newest obtained position \( y_n^t \) act as the input for the fitness function to evaluate whether a value is the current best result and updated with the BA rules [13]. To evaluate the proposed model, MBA, four test functions and four comparison algorithm are employed. The maximum generation and population size of MBA are set as 10000 and 100, respectively. Meanwhile, the iteration number of inner loop of MBA is set as 5 to avoid results falling into local optimums.

| Fun. # | Test function          | Tolerance value |          |          |          |          |
|-------|-----------------------|-----------------|----------|----------|----------|----------|
|       |                       | Min             | Max      | Mean     | Std.     |          |
| 1     | Beale function        | FPA             | 6.4625e-27 | 5.0827e-20 | 9.6821e-24 | 4.6967e-24 |
|       |                       | SAPSO           | 4.0952e-10 | 2.0843   | 5.0063e-04 | 6.8036e-05 |
|       |                       | BA              | 6.6396e-12 | 0.4903   | 4.0854e-03 | 6.0646e-03 |
|       |                       | SAGA            | 5.6008e-14 | 4.0841e-10 | 7.2084e-12 | 9.0482e-11 |
|       |                       | MBA             | 0         | 0        | 0        | 0        |
| 2     | Dixon and Price’s function | FPA             | 3.8403e-17 | 4.5093e-11 | 8.0936e-15 | 7.9348e-15 |
|       |                       | SAPSO           | 4.3093e-16 | 7.2552e-10 | 8.3443e-14 | 6.4883e-14 |
|       |                       | BA              | 1.9903e-14 | 1.5323e-10 | 3.5563e-12 | 8.3532e-13 |
|       |                       | SAGA            | 5.4753e-14 | 6.3448e-08 | 7.8347e-11 | 6.4384e-11 |
|       |                       | MBA             | 5.0835e-34 | 6.4773e-24 | 3.9845e-29 | 4.4885e-30 |
| 3     | Griewank’s function   | FPA             | 6.3672e-08 | 9.5549e-05 | 6.5645e-06 | 4.4902e-05 |
|       |                       | SAPSO           | 3.4803e-14 | 5.0943e-08 | 8.5461e-11 | 4.6064e-11 |
|       |                       | BA              | 1.8846e-12 | 3.0849e-03 | 4.6783e-06 | 9.8735e-07 |
|       |                       | SAGA            | 5.9834e-16 | 9.5314e-13 | 1.6804e-14 | 6.5495e-13 |
| 4     | Rotated Hyper-Ellipsoid function | FPA             | 7.5485e-67 | 6.4553e-62 | 6.4654e-64 | 2.7382e-63 |
|       |                       | SAPSO           | 6.6733e-29 | 1.3638e-24 | 5.8601e-16 | 1.9253e-15 |
|       |                       | BA              | 5.7445e-24 | 7.5633e-20 | 8.4552e-22 | 9.8067e-21 |
|       |                       | SAGA            | 1.4453e-18 | 1.6456e-12 | 3.5648e-16 | 3.1249e-15 |
As the test results shown in Table 1, MBA achieves the best results among five algorithms. According to the comparison between original BA, MBA enhanced the local convergence ability largely.

3. MBA-ELM

ELM is a powerful and fast learning algorithm based on the modification of the traditional single-hidden layer feed-forward neural networks (SLFNN). It has been proved that the ELM network has better accuracy performance than the ANNs and the support vector machine.

Here, the network weights \( \omega \) between the input layer and the hidden layer are defined as
\[
\omega = \begin{bmatrix}
\omega_{11} & \omega_{12} & \cdots & \omega_{1n} \\
\omega_{21} & \omega_{22} & \cdots & \omega_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\omega_{l1} & \omega_{l2} & \cdots & \omega_{ln}
\end{bmatrix}_{l \times n}
\]
(2)

where \( l \) is the number of the hidden neurons and \( n \) is the number of the input neurons. The network weights \( \beta \) between the hidden layer and the output layer are defined as
\[
\beta = \begin{bmatrix}
\beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\
\beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
\beta_{m1} & \beta_{m2} & \cdots & \beta_{mm}
\end{bmatrix}_{m \times l}
\]
(3)

where \( m \) is the number of the output neurons. And the thresholds \( b \) of the hidden layer are defined as
\[
b = \left[ b_1, b_2, \ldots, b_l \right]_{l \times 1}
\]
(4)

Before the training process, the initial values of \( \omega, \beta \) and \( b \) are set randomly. However the final results of ELM are strongly dependent on the initial weights and thresholds, and the bad initial weights may lead to a slow convergence of optimal value in ELM. To avoid wasting more training costs, an optimization method is proposed to determine the initial weights and thresholds based on MBA.

There are two processes consisted in MBA-ELM: optimization of the values of \( \omega, \beta \) and \( b \) and ELM training. The structure of MBA-ELM is shown in Figure 1.

Figure 1. Structure of MBA-ELM

To set the initial values of weights and thresholds of ELM to avoid wasting training costs, MBA is employed to optimize \( \omega, \beta \) and \( b \). The fitness function is expressed as follows.
\[
f(\omega, \beta, b) = \min \left[ T - \hat{T}(\omega, \beta, b) \right]
\]
(5)

Here, the elements in \( T \) are the actual values of the wind speed data.
and $\hat{T}(\omega, \beta, b)$, which represents the outputs of the ELM network is defined as:

$$\hat{T}(\omega, \beta, b) = \left[\hat{t}_1, \hat{t}_2, \hat{t}_3, \ldots, \hat{t}_p\right]$$

and the forecasting results of ELM and MBA are shown in Table 2. The detailed multistep promoting percentages of the hybrid models of the four sites are shown in Table 3.

4. Forecasting Experiment

Three data sites in Penglai region have been selected. Meanwhile, considering randomness factors and ensuring that the final results are reliable and independent of the initial weights, we carry out each experiment 50 times and then take the average value. The input layer of all the ANNs is constructed with four neurons. To determine the node number of the hidden layer, with the usage of Hecht–Nelson method, we denote the node number of the hidden layer is $2n + 1$ when the node number of the input layer is $n$.

We choose three metric parameters to compare whether the proposed model optimized the initial ELM network: the MAE (mean absolute error), the MAPE (mean absolute percentage error) and the MSE (mean square error). Additionally, to obtain the detailed promoting percentages when comparing two forecasting models, MAE percentage error indexes are denoted as follows, so do $\xi_{MAE}$, $\xi_{MAPE}$.

$$\xi_{MAE} = \frac{MAE_t - MAE_i}{MAE_i} \times 100\%$$

To conduct the experiment, here we chose 1500 history data for training and 500 data for testing. The multistep forecasting results of ELM and MBA-ELM are shown in Table 2. The detailed multistep promoting percentages of the hybrid models of the four sites are shown in Table 3.

| Site 1  | MAE    | MSE    | MAPE    | 1-step | 2-step | 3-step |
|--------|--------|--------|---------|--------|--------|--------|
|        | 0.138  | 0.233  | 0.322   | 0.096  | 0.180  | 0.298  |
| Site 2  | 0.159  | 0.245  | 0.353   | 0.109  | 0.204  | 0.277  |
| Site 3  | 0.130  | 0.280  | 0.397   | 0.122  | 0.221  | 0.302  |
MAPE  1.687  3.423  5.126  1.516  3.120  4.096  
MAE    0.175  0.244  0.386  0.132  0.243  0.319  
MSE    0.026  0.108  0.173  0.025  0.075  0.129  
MAPE   1.759  3.298  5.205  1.654  3.181  4.239  
Site 4

Table 2 indicates the following:
(1) For four sites, the forecasting results of MBA-ELM are more accurate than ELM from one-step forecasting to three-step forecasting.
(2) One-step forecasting of ELM and MBA-ELM is the most accurate among one-step to three-step forecasting.

Table 3. Improvement Percentages Between ELM and MBA-ELM

| Site   | ELM vs MBA-ELM | 1-step | 2-step | 3-step |
|--------|----------------|--------|--------|--------|
| Site 1 | zm, zm, zm    | 30.435 | 22.747 | 7.453  |
|        | zm, zm, zm    | 76.364 | 25.806 | 40.769 |
|        | zm, zm, zm    | 14.197 | 11.075 | 8.075  |
| Site 2 | zm, zm, zm    | 31.447 | 16.735 | 21.530 |
|        | zm, zm, zm    | 53.191 | 29.114 | 38.788 |
|        | zm, zm, zm    | 4.904  | 19.832 | 20.647 |
| Site 3 | zm, zm, zm    | 6.154  | 21.071 | 23.929 |
|        | zm, zm, zm    | 32.558 | 42.241 | 33.758 |
|        | zm, zm, zm    | 10.136 | 8.852  | 20.094 |
| Site 4 | zm, zm, zm    | 24.571 | 0.410  | 17.358 |
|        | zm, zm, zm    | 3.846  | 30.556 | 25.434 |
|        | zm, zm, zm    | 5.969  | 3.548  | 18.559 |

Table 3 illustrates the following:
(1) In the one-step predictions, the MAPE-promoted percentage from four sites of MBA-ELM by ELM is 14.197%, 4.904%, 10.136% and 5.969%, respectively.
(2) In the two-step predictions, the MAPE-promoted percentage from four sites of MBA-ELM by ELM is 11.075%, 19.832%, 8.852% and 3.548%, respectively.
(3) In the three-step predictions, the MAPE-promoted percentage from four sites of MBA-ELM by ELM is 8.075%, 20.647%, 20.094% and 18.559%, respectively.

Note that by comparing the proposed hybrid model, MBA-ELM and ELM, the MBA-ELM hybrid model has the most accurate forecasting results. That indicates MBA could improve the forecasting performance of ELM.

Forecasting valid degree (FVD) is defined to evaluate the forecasting accuracy of the proposed hybrid model MBA-ELM and ELM. A more accurate forecasting model leads to a larger FVD value. The results presented in Table 4 show that the FVD value for the MBA-ELM model is larger than ELM.

Table 4. Results for FVD

| Model    | FVD        | 1-step | 2-step | 3-step |
|----------|------------|--------|--------|--------|
| ELM      | 98.2513    | 96.5415 | 94.7943 |
| MBA-ELM  | 98.4050    | 96.9243 | 95.6683 |

5. Conclusion
In this paper, a hybrid model based on modified bat algorithm is proposed for multi-step wind speed forecasting. Meanwhile, to improve the performance of ELM, a new improved BA, modified bat algorithm based on conjugate gradient method, was proposed by us to optimize the initial weights and thresholds of ELM. The observed numerical results can be summed up into the following conclusion.
First, according to the four test functions results, the performance of MBA is better than BA; second, MBA improves the performance of ELM with more accurate results obtained. Therefore, the proposed MBA-ELM model performs well in forecasting wind velocity.

Further research should be conducted to improve the predicting accuracy as the forecasting step increases. Besides, other numerical optimization algorithms should be proposed to improve the performance of ANN and make the comparison between differential methods.

Acknowledgments
Special thanks shall be paid to Liye Xiao (University of Electronic Science and Technology of China (UESTC)), who gives the authors the guidance and advice on this paper’s experiment and writing parts.

References
[1] Chen, H. L., Yang, B., Wang, G., Liu, J., Xu, X., Wang, S. J., & Liu, D. Y. (2011). A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. Knowledge-Based Systems, 24(8), 1348-1359.
[2] Wu, Y. K., & Hong, J. S. (2007, July). A literature review of wind forecasting technology in the world. In Power Tech, 2007 IEEE Lausanne (pp. 504-509). IEEE.
[3] Ranjan, J., Goyal, D. P., & Ahson, S. I. (2008). Data mining techniques for better decisions in human resource management systems. International Journal of Business Information Systems, 3(5), 464-481.
[4] Li, H., Wang, J., Lu, H., & Guo, Z. (2018). Research and application of a combined model based on variable weight for short term wind speed forecasting. Renewable Energy, 116, 669-684.
[5] Cassola, F., Villa, L., Burlando, M., & Ratto, C. F. (2008). Downscaling NWP wind fields by means of a diagnostic mass-consistent model. National Wind Engineering Congress.
[6] Liu, H., Tian, H. Q., Liang, X. F., & Li, Y. F. (2015). Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Applied Energy, 157, 183-194.
[7] Alboyaci, B., & Dursun, B. (2008). Electricity restructuring in Turkey and the share of wind energy production. Renewable Energy, 33(11), 2499-2505.
[8] Wang, S., Zhang, N., Wu, L., & Wang, Y. (2016). Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. Renewable Energy, 94, 629-636.
[9] Yang, X. S., & Hossein Gandomi, A. (2012). Bat algorithm: a novel approach for global engineering optimization. Engineering Computations, 29(5), 464-483.
[10] Qu, Z., Zhang, K., Mao, W., Wang, J., Liu, C., & Zhang, W. (2017). Research and application of ensemble forecasting based on a novel multi-objective optimization algorithm for wind-speed forecasting. Energy Conversion and Management, 154, 440-454.
[11] Zhang, C., Zhou, J., Li, C., Fu, W., & Peng, T. (2017). A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. Energy Conversion and Management, 143, 360-376.
[12] Xiao, L., Qian, F., & Shao, W. (2017). Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm. Energy Conversion and Management, 143, 410-430.
[13] Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: theory and applications. Neurocomputing, 70(1-3), 489-501.