A New Method for Axis Adjustment of the Hydro-generator Unit Using Machine Learning

Jie Cao (✉ caojie@neepu.edu.cn)

Northeast Electric Power University

Yang Li

Northeast Electric Power University

Zhaoyang Qu

Northeast Electric Power University

Yunchang Dong

Northeast Electric Power University

Yaowei Liu

State Grid Jilin Electric Power Co, LTD. ,Changchun, China

Ruxuan Zhang

Northeast Electric Power University

Research Article

Keywords:

Posted Date: February 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1278166/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A new method for axis adjustment of the hydro-generator unit using machine learning

Jie Cao^{1,2*}, Yang Li¹, Zhaoyang Qu¹, Yunchang Dong³, Yaowei Liu⁴, Ruxuan Zhang¹

¹School of Computer Science, Northeast Electric Power University, Jilin, China

²Guangdong Atv Academy For Performing Arts

³School of Electrical Engineering, Northeast Electric Power University, Jilin, China

⁴State Grid Jilin Electric Power Co., LTD., Changchun, China

The power quality and efficiency of the hydro-power station depend on the stable operation of the hydro-generator unit, which needs to continue to operate and it is prone to axis failure. Therefore, to adopt effective axis adjustment technology to eliminate faults. This paper proposes a new method for axis adjustment of hydro-generator unit based on an improved grey prediction model and swarms intelligence optimization neural network. First of all, it proposes a sequence acceleration translation and mean value transformation method, which is used to pre-process the axis net total swing sequence that exhibits oscillating fluctuations. It uses e_1 and e_2 factor transformation to establish an improved axis net total swing gray prediction model. Then, the advanced flamingo search algorithm is used to search the maximum value of the sine function of the net total pendulum of the axis, and the axis adjustment orientation is obtained. This method solves the problem that GM(1, 1) can only be predicted by monotone sequence in the past and the problem that the search algorithm is easy to fall into local optimum, effectively improves the calculation efficiency of axis and shorts the search time. Simulation examples show that the proposed method can significantly improve accuracy of axis adjustment. This method greatly improves the efficiency of azimuth search for axis adjustment.

Hydro-generator unit (HGU) is a part of the key equipment of hydro-power station¹⁻². State of the reliable measurement of the trend of hydroelectric generating set axis for unit security, it promotes the power system stability of great significance³. In practical engineering, axis adjustment is the most important work in the later period of unit installation. Overhaul of the unit must also go through axis adjustment inspection. The quality of the unit axis comprehensively reflects the quality of installation and maintenance⁴. Axis adjustment central axis of the net total swing and maximum net total swing degrees is an important parameter of axis adjustment calculation, so the degree of the axis of the hydroelectric generating set net total swing and prediction as well as to the axis of maximum net total swing adjustment bearing search, to guarantee the safety of the hydro-generator unit and reduce economic loss of the hydro-power station is of great significance⁵. The calculation of net total swing and relative swing in axis adjustment of the hydro-generator unit is still mainly by manual calculation. The Axis calculation process is complicated, it is involving the measurement data of upper lead, lower lead, water lead, and other parts, which need the dimensions of upper lead to lower lead, upper lead to water lead, thrust head diameter, and so on. Therefore, the efficient and accurate prediction method of axis net total swing is conducive to reducing the number of axis measurements, and it is also the basis and prerequisite for realizing axis measurement⁶. There are four typical prediction algorithms for the prediction of net total swing and maximum net total swing of hydro-generator axis in machine learning methods: the prediction method using BP neural network (BP); Prediction method using support vector machine (SVM); Prediction method using random forest (RF); Prediction method using grey model (GM). There are two main methods of axis adjustment azimuth search: traditional manual processing and swarm intelligence algorithm. At present, machine learning has been widely used in many fields, but not many applications have been applied to the axis adjustment of the hydro-generator unit⁷. However, for axis adjustment orientation, machine learning can solve the problem that axis adjustment orientation is difficult to search by using algorithms⁸⁻⁹.

To solve the above problems, this paper proposes a new method for axis adjustment of the hydro-generator unit using machine learning. Its main contributions include:

1. A prediction method of axis net total swing using the improved GM(1, 1) is proposed. Through accelerating translation and meaning transformation methods, the net total swing sequence shows oscillatory fluctuations are pre-processed to weaken the series volatility. This model solves the problem that GM(1, 1) can only be predicted by monotone sequence. Then, the improved GM(1, 1) algorithm can be used to predict axis net total swing data and improve the axis measurement processing efficiency.
2. An optimization method of axis adjustment orientation search using flamingo search algorithm is proposed to accurately search axis adjustment orientation. This model solves the problem that traditional search algorithm is easy to fall into local optimal solution.

3. This paper proposed method with the traditional manual axis processing, and compare it with the representative machine learning algorithm. The performance evaluation shows that can significantly improve the search of axis adjustment orientation by the model.

In the area of axial net total swing prediction, a classifier combining rough sets and the support vector machine was proposed in the literature [11], it was applied to fault diagnosis of HGU. An HPG vibration trend measurement model was proposed in the literature [12], which is based on optimal variational mode decomposition (OVMD) and chaotic sine-cosine algorithm optimization (CSCA) and improved least squares support vector machine (LSSVM). The literature [13] proposed a Guassion regression process (GPR) based early fault prediction method for HPG, which effectively tracked the change process of HPG operation status and monitors the abnormal operation status of HPGs in advance. Literature [14] proposed a least squares support vector machine (LS-SVM) method for predicting the stability parameters of mixed-flow hydro turbine units. The literature [15] developed a time series prediction model ARIMAX (1, 1, 1) for hydroelectric power generation in ecuador to predict the monthly output for up to one year. The literature [16-18] proposed a grey prediction model (GM(1, 1)) for predicting the performance degradation trend of hydroelectric units. Although the above methods improved the efficiency of the axial measurement under certain conditions, the selection of the neural network needs to consider the size of the data set as well as the setting of the parameters, so there are inevitable errors.

In terms of axial adjustment orientation search, literature [19-20] proposed an improved adaptation evaluation method, it put forward the concept of stability and instability parameter sets, and it applied to the stability search strategy, which effectively improved the local search capability of the model algorithm and it ensured the stability and accuracy of the hydro-generator set model. The literature [21-24] proposed a prediction method based on radial basis function (RBF) interpolation, empirical modal decomposition (EMD), approximate entropy, artificial neural network, and grey theory for the degradation of hydro-generator unit characteristic parameters. A nonlinear prediction model for the degradation trend of hydro generator unit operating parameters was proposed in the literature [25-29]. The model was based on radial basis function interpolation, wavelet transform, maximum lyapunov exponent prediction method, and gray prediction model (GM(1, 1) method). The literature [30] put forward the dynamic alarm curve model, established GA-BP neural network model, and it obtained the vertical vibration of head covering by training and outputting the alarm dynamic curve. Although the above methods can predict the axis adjustment amount at a certain moment, most of them belong to single-step prediction, and when it is necessary to know the changing trend of the axis adjustment amount of the hydro-generator unit earlier, the single-step prediction model needs to be extended to a multi-step prediction using rolling prediction [31-33], which time the prediction accuracy will be reduced due to the continuous superposition of errors each time, which cannot meet the actual need of hydro-generator set maintenance project. The machine learning algorithm is an important method to solve prediction and optimization. Based on the improved grey model, the data series with oscillating wave phenomena are pre-processed by accelerating translation and meaning transformation method to weaken the series volatility [34-36], and then the prediction function is obtained by factor transformation. The flamingo algorithm is a new swarm intelligence optimization algorithm based on the migration and foraging behavior of flamingos [37-39]. It can increase the ability of searching and develop search space, ensure a good balance between search and development, and it can effectively solve nonlinear optimization problems. The improved RBF neural network uses particle swarm optimization to determine the center of the RBF neural network [40-42], it controls the optimization speed by the inertia weight factor, and it takes the maximum net total swing of the maintenance data of the hydro-generator unit as the input vector of the neural network to predict the real axis adjustment amount [43-45].

Methods

In this paper, the net total swing at the axis adjustment point is used as the state quantity to represent the operating state of the hydro-generator unit. The improved grey model is used to predict the net total swing at the axis adjustment point in the future, and the predicted values of the net total swing of the axis at eight state points around the big axis are obtained respectively. Next, the sine function fitting of the net total swing value of the eight-point axis is carried out to obtain the sine function, and then, on the basis of obtaining the sine function of the net total swing. Then, using the obtained net total swing sine function, the flamingo search algorithm is used to search the maximum value of the sine function of net total swing at the axis adjustment, and the corresponding axis adjustment orientation is obtained.

The determination of axis adjustment orientation and axis adjustment amount is also realized. The general framework is shown in Figure 1.

Figure 1. Framework diagram of the axial alignment model

In this section, we review the relevant data of the hydro-generator unit and observing the axis maintenance data set, it is found that the axis net total swing sequence shows a fluctuation phenomenon. To solve the problem that the axis net total swing has oscillation fluctuation, this paper uses accelerated translation and weighted mean transformation to preprocess the axis net total pendulum with fluctuation phenomenon, and then improved GM(1, 1) modeling after meeting the monotone condition.

Net total swing prediction based on improved GM(1, 1) model

GM(1, 1) model. The basic idea of the traditional GM(1, 1) model: To facilitate mathematical modeling, the original net total swing sequence is accumulated once. Since the sequence after accumulation has an exponential growth trend, the approximate first-order differential equation is used to build the model. Finally, the prediction sequence is generated by the accumulation of the modeling sequence to complete the prediction of the development trend of the original net total swing sequence.

The specific modeling process for the traditional GM(1, 1) is as follows.

Let the original net total swing sequence be: $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$, and it performs a single accumulation on this sequence to generate. As equation (1).

$$X^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), K = 1, 2, \dots, n \quad (1)$$

Generate a net total swing sequence with exponential regularity as $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$.

$X^{(1)}$ approximates the sequence as a first-order differential (2) equation's solution.

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (2)$$

Where: a is the development coefficient of the model; b is the amount of grey action.

Denote the parameter $A = [a, b]^T$, and it uses the least-squares method to find A as

$$A = (B^T B)^{-1} B^T Y \quad (3)$$

Where:

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \text{MM} \\ -\frac{1}{2}[x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{bmatrix}$$

$$Y = \{x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)\}^T$$

Find the values of a , b and substitute them into equation (2) to calculate.

$$\hat{x}^{(1)}(k+1) = [x^{(0)}(1) - \frac{b}{a}]e^{-ak} + \frac{b}{a} \quad (4)$$

The cumulative reduction from equation (4) yields the net total swing prediction function as

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = (1 - e^a)[x^{(0)}(1) - \frac{b}{a}]e^{-ak} \quad (5)$$

In the traditional GM(1, 1) modeling, because the net total swing of the original axis sequence is not very regular, it is necessary to carry out cumulative changes and then it uses equation (2) to build a mathematical model. This modeling method regardless of whether the axis net total swing original sequence oscillates or not, the cumulatively generated sequence will change monotonically, and the reduced sequence also shows the same change trend. When the net total swing of the original sequence is monotonic, the prediction accuracy is better, but when the original sequence oscillates and fluctuates, the prediction accuracy is not ideal because the reduced sequence changes monotonically and the original swing sequence cannot be fitted accurately.

If the axis net total swing oscillation wave sequence is transformed mathematically to make it have a monotone trend, the mathematical model is established by GM(1, 1), finally, the reduction function is calculated, and then the mathematical inverse transformation is carried out to obtain the axis net total swing prediction sequence, it will be a good solution to the problem that the traditional GM(1, 1) model does not have high prediction accuracy for the oscillating sequence. In this paper, the accelerated translation and meaning transformation method are combined to pre-process the series with oscillatory fluctuation phenomenon, to weaken the fluctuation of the series.

$X = \{x(1), x(2), \dots, x(n)\}$ is the axis net total swing original sequence, and if there exists $k, k \in [1, 2, \dots, n-1]$, such that $x(k+1) - x(k) > 0, x(k+1) - x(k) < 0$, then it is said to X be a sequence of random swings. As equation (6) and (7).

$$M = \max\{x(k) \mid k = 1, 2, \dots, n\} \quad (6)$$

$$m = \min\{x(k) \mid k = 1, 2, \dots, n\} \quad (7)$$

$M-m$ is the amplitude of the sequence X , denoted as T .

Define accelerated translation transformation: In order to weaken the volatility of the axis net total swing of the original oscillatory sequence. $XE_1 = \{x(1)e_1, x(2)e_1, \dots, x(n)e_1\}$, Eq:

$$x(k)e_1 = x(k) + (k-1)T, k = 1, 2, \dots, n \quad (8)$$

It is called e_1 the accelerated translation transformation factor, and the monotonic trend of the original swing sequence after this transformation can be proved by simple mathematical operations.

Define the weighted mean transformation: the axis net total swing sequence after e_1 factor processing has shown a monotonic trend and it has modeling conditions. To fit the axis net total swing of the original sequence more accurately, the weighted mean transformation is introduced to perform a quadratic transformation on the factorized sequence, and the transformed sequence is defined.

$XE_2 = \{x(1)e_1e_2, x(2)e_1e_2, \dots, x(n)e_1e_2\}$, Eq:

$$x(k)e_1e_2 = (\sum_{i=1}^k x(i)e_1) / k \quad (9)$$

It is called e_2 the weighted mean-transformed factor. Through simple mathematical calculation, it can be proved that the axis net total swing sequence after e_2 factor transformation keeps the monotone characteristics of the axis net total swing original sequence, which can make the axis net total swing sequence smoother.

Improved GM(1, 1) modeling of oscillatory sequences. Let the axis net total swing of the original swing sequence be

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \Lambda, x^{(0)}(n)\} \quad x^{(0)}(k) > 0, k = 1, 2, \Lambda, n$$

The improved GM(1, 1) modeling processes are as follows.

(1) $X^{(0)}$ Perform an e_1 factor transformation on the axis net total swing of original sequence as $Y^{(0)} = \{y^{(0)}(1), y^{(0)}(2), \Lambda, y^{(0)}(n)\}$

(2) The sequence $Y^{(0)}$ is e_2 factor-transformed to

$$Z^{(0)} = \{z^{(0)}(1), z^{(0)}(2), \Lambda, z^{(0)}(n)\}$$

(3) Perform a summation of the sequence $Z^{(0)}$.

$$Z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), \Lambda, z^{(1)}(n)\}$$

(4) Modeling $Z^{(1)}$'s GM(1, 1) differential equation for the sequence. As equation (10).

$$\frac{dz^{(1)}}{dt} + az^{(1)} = b \quad (10)$$

a, b parameters by least squares. As equation (11).

$$\begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T P \quad (11)$$

It is determined, among other things.

$$B = \begin{bmatrix} -\frac{1}{2}[z^{(1)}(1) + z^{(1)}(2)] & 1 \\ -\frac{1}{2}[z^{(1)}(2) + z^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}[z^{(1)}(n-1) + z^{(1)}(n)] & 1 \end{bmatrix}$$

$$P = [z^{(0)}(2), z^{(0)}(3), \Lambda, z^{(0)}(n)]^T$$

(5) The response function of the GM(1, 1) differential equation.

$$z^{(1)}(k+1) = [z^{(0)}(1) - \frac{b}{a}]e^{-ak} + \frac{b}{a} \quad (12)$$

(6) The axis net total swing prediction function is obtained by one cumulative reduction $z^{(0)}$.

$$z^{(0)}(k+1) = z^{(1)}(k+1) - z^{(1)}(k) = (1 - e^a)[z^{(0)}(1) - \frac{b}{a}]e^{-ak} \quad (13)$$

Which: $k = 1, 2, \Lambda, n; z^{(0)}(1) = z^{(0)}(1)$.

(7) For $Z^{(0)}$ use e_2 factor inverses transformation reduction to attain. As equation (14).

$$y^{(0)}(k) = k z^{(0)}(k) - \sum_{i=1}^{k-1} y^{(0)}(i), k = 2, 3, \Lambda, n \quad (14)$$

Where $y^{(0)}(1) = y^{(0)}(1)$, and

$$y^{(0)}(k+1) = (k+1)z^{(0)}(k+1) - k z^{(0)}(k) \quad (15)$$

(8) The inverse transformation of the axis net total swing of the original sequence by e_1 factor pair $Y^{(0)}$ yields a prediction function.

As equation (16).

$$x^{(0)}(k+1) = y^{(0)}(k+1) - kT = (k+1)z^{(0)}(k+1) - k z^{(0)}(k) - kT \quad (16)$$

Which: $k = 1, 2, \Lambda, n; x^{(0)}(1) = x^{(0)}(1)$.

The above improved steps of GM(1, 1) and the formula after factor transformation are shown in Figure 2.

Figure 2. Framework diagram of the improved GM algorithm

Adjusted orientation finding model based on the flamingo search algorithm

Obtained by the improved GM(1, 1) to predict the axis of the net total degree of the pendulum, the next it needs to be clean all degrees of sine function fitting and establish a large shaft bearing as the abscissa, net total swing as the ordinate axis net total swing degrees after sine fitting function, using the flamingo algorithm to the maximum net total swing of sine function optimization. Finally, the independent variable corresponding to the maximum value of sine function of axis net total pendulum is the axis adjustment azimuth.

Flamingo has the global search and local development capabilities needed for optimization algorithms. It has good results for single-peak and multi-peak function search. Therefore, it is suitable for searching the maximum value of the sine function of net total swing.

The two main behavioral traits of flamingos are foraging and migratory behavior. Flamingos primarily inhabit areas where food is plentiful. After a period of extensive foraging, the flamingo population migrates when the food in the area decreases to a level that does not satisfy the population. Corresponding foraging and migration models are developed.

The main optimization ideas of the FSA model are as follows.

Foraging behavior

1. Communicative behavior

Flamingos with the most food in the group spread their position information by calling other flamingos, and it influences the position changes of other flamingos in the group. Theoretically, flamingos cannot know where the most food is in an area. However, this does not mean that the algorithm cannot find a globally optimal solution, because the algorithm is a program, you can't tell the end condition of the program when you set it.

FSA is an algorithm that simulates flamingos searching for the best solution in a search area based on the limited available information. In this paper, we assume that the flamingo with the largest amount of food in the j th dimension is xb_j .

2. Beak scanning behavior

When a flamingo's beak is poured over water, it acts like a large sieve, and it is sucking the water and then quickly filtering it out, a foraging pattern influenced by the abundance of food in the area. If there is more food in the area swept by the flamingo's bill, this encourages the flamingo to scan the area more carefully and the flamingo's neck slowly stretches out, it is causing the beak's scan radius to increase. The likelihood of scanning the area for food will also increase, and a model of the flamingo's beak scanning behavior is shown in Figure 3.

Figure 3. Beak scanning behavior

The closer a flamingo is to the most abundant food in the population, the greater the likelihood that food is abundant in that area. In this paper, we simulate the beak-scanning behavior of flamingos. Let the position of the i th flamingo on the j th dimension of the flamingo population is x_{ij} , it is necessary to consider the variability of flamingo individual selection in nature and the suddenness of specific environment, which will affect the foraging behavior of flamingo. Otherwise, there are errors in the foraging behavior and information transmitted by flamingos. To model this error, a standard normal random distribution is introduced, in which the flamingo beak scan has a high probability of aligning with the position where the food is most abundant. However, there is also a small probability of error based on this information.

The maximum distance of flamingo beak scans in foraging behavior can be quantified as $|G_1 \times xb_j + \varepsilon_2 \times x_{ij}|$, where ε_2 is a random number of -1 or 1. The maximum distance is primarily intended to increase the search range of the flamingo beak scan in foraging mode, where G_1 is a random number that fits a standard normal distribution. To model the scanning range of the flamingo in beak scanning behavior, the normal distribution is again introduced, and its variation curve approximates the variation of the flamingo's beak scanning range as $G_2 \times |G_1 \times xb_j + \varepsilon_2 \times x_{ij}|$, where G_2 is a random number obeying a normal distribution.

3. Bipedal mobile behavior

A model of the foot movement behavior of flamingos is shown in Figure 4. When flamingos forage for food, as they scan with their beaks, their talons move towards the most abundant food in the flamingo population. The distance traveled can be quantified by assuming that the most abundant food place in the population is $\varepsilon_1 \times xb_j$, ε_1 as a random number of -1 or 1, it mainly to increase the search range of flamingos foraging and to quantify individual selection differences.

Figure 4. Bipedal mobile behavior

In summary, the moving step of the flamingo foraging in the t iteration is the scanning range of the flamingo beak, and then the moving distance of its feet is added, as shown in (17).

$$b_{ij}^t = \varepsilon_1 \times xb_j^t + G_2 \times |G_1 \times xb_j^t + \varepsilon_2 \times x_{ij}^t| \quad (17)$$

The equation to update the location of flamingo foraging behavior is

$$x_{ij}^{t+1} = (x_{ij}^t + \varepsilon_1 \times xb_j^t + G_2 \times |G_1 \times xb_j^t + \varepsilon_2 \times x_{ij}^t|) / K \quad (18)$$

In (18), x_{ij}^{t+1} denotes the position x_{ij}^t of the i th flamingo in the j th dimension of the population in the $(t + 1)$ th iteration, x denotes the position of the i th flamingo in the j th dimension of the flamingo population in the t th iteration, i.e., the position of the flamingo foot. xb_j^t denotes the j th dimensional position in the population of the best adapted flamingo in t iterations. $k = K(n)$ is the diffusion factor, which is a random number with a cardinal distribution of n degrees of freedom. It is used to increase the size of the flamingo's foraging range and to simulate the chances of individual selection in nature, it is increasing global meritocracy. $G_1 = N(0,1)$ and $G_2 = N(0,1)$ are random numbers that follow a standard normal distribution, ε_1 and ε_2 are randomized by -1 or 1.

Migration behavior

When food is scarce in the current foraging area, the flamingo population migrates to the next more abundant food area. Assuming that the location of the food rich area in dimension j th is xb_j , the migration equation for the flamingo population is as follows.

$$x_{ij}^{t+1} = x_{ij}^t + \omega \times (xb_j^t - x_{ij}^t) \quad (19)$$

In (19), x_{ij}^{t+1} denotes the position of the i th flamingo in the j th dimension of the population in $t+1$ iterations, x_{ij}^t denotes the position of the i th flamingo in the j th dimension of the flamingo population in t iterations, i.e., the position of the flamingo foot. xb_j^t denotes the j th dimensional position of the adapted flamingo in the population in t iterations. $\omega = N(0, N)$ is a random number with N degrees of freedom, it is used to increase the search space during the migration of flamingos, and it is also used to simulate the randomness of individual behavior of flamingos during specific migration.

The algorithmic flow of FSA

Step1: Initialize the axis net total swing function population, with the population set to P and the maximum number of iterations to $Iter_{Max}$, and the proportion of migrating flamingos in the first part MP_b .

Step2: The number of foraging flamingos in the i th iteration of the flamingo population update is $MP_i = rand[0,1] \times P \times (1 - MP_b)$. The number of migrating flamingos in the first part of this iteration is $MP_0 = MP_b \times P$. The number of migrating flamingos in the second part of this iteration is $MP_i = P - MP_0 - MP_i$, and the fitness values of individual flamingos are obtained, and the flamingo population is ranked according to the fitness values of the individuals. The MP_b former flamingos with low fitness and the former flamingos MP_i with high fitness are considered migratory flamingos, while the others are considered foraging flamingos.

Step3: Update migrating flamingos according to (19) and foraging flamingos according to (18).

Step4: Check for out-of-bounds flamingos.

Step5: Reach the maximum number of iterations and go to Step6; otherwise, execute Step2.

Step6: Output axis net total swing function's the optimal solution and the optimal value. The flow chart of FSA is shown in Figure 5.

Figure 5. Flow chart of FSA algorithm

Results and discussion

Datasets. This paper is based on the maintenance data of a hydraulic power maintenance company from 2015 to the 2020 year. Due to the small number of maintenance visits of the disassembled generator set every year, there are 200 pieces of maintenance data in this data set, among which each piece of data has 4 kinds of characteristic data, and it is including absolute swing, relative swing, net total swing and maximum net total swing. Because this type of unit is commonly used by hydro-electric power maintenance companies, this data set has become an important reference data set. In this paper, a prediction study is conducted on the 1-80 times data of the maintenance data set, which consists of eight measuring points in a circle of the large axis and a total of 2400 items. Among them, there are 640 pieces of net total swing data. Maintenance workers mainly use the net total swing as a reference standard to carry out work. Therefore, net total swing and maximum net total swing are screened separately as data sets. The data of the first 70 items are the training set, and the data of the last 10 items are the test set. 71-80 times of data are used for research in this paper. Table 1 is the data of the net total pendulum.

Serial number	Measurement Point I (mm)	Measurement Point II (mm)	Measurement point III (mm)	Measurement Point IV (mm)	Measurement Point V (mm)	Measurement point VI (mm)	Measurement point VII (mm)	Measurement point VIII (mm)
1	25	37	49	8	-25	-37	-49	-8
2	14	-4	-20	-26	-14	4	20	26
3	12	-4	-20	-24	-12	4	20	24
4	16	20	25	3	-16	-20	-25	-3
5	19	23	14	-2	-19	-23	-14	2
6	14	21	16	-3	-14	-21	-16	3

7	16	20	16	-4	-16	-20	-16	4
8	14	-7	-24	-26	-14	7	24	26
9	-17	-19	-11	7	17	19	11	-7
10	6	3	-2	-4	-6	-3	2	4

Table 1. Net total swing of measurement points (71-80 times)

The data set is obtained by using the Marr scale and measurement software. The data measured by the Marr scale is an absolute pendulum, and then the net total pendulum, relative total pendulum, and maximum net total pendulum are obtained by a series of axis adjustment formulas.

To verify the accuracy and validity of the improved GM(1, 1) model established in this study, historical data at the suspended second guide flange are selected and predictions are obtained 1-80 times, and GM(1, 1), BP, SVM, and RF are selected for comparison.

The pendulum trend of the net total swing of the axis of the eight measurement points can be observed in Figure 6, from which it can be found that the data of each measurement point shows oscillatory nature, it is not a single increasing or decreasing trend, so the traditional GM(1, 1) is not very applicable to it, the oscillation trend needs to be weakened.

Figure 6. Trend of net full swing at eight measurement points on the axis

In this experiment, the axis net total swing sequence with oscillating wave phenomenon is processed by accelerating translation and meaning transformation method to weaken the net total swing oscillation sequence. GM(1, 1) is improved by using this method, and the net total swing is taken as a predictive variable. At the same time, GM(1, 1), BP, SVM, and RF models are selected for comparison, and the net total swing of 71-80 times axis is predicted respectively, with the first 70 times as the training value and 71-80 times as the predicted value. The actual value; predicted value, and relative error of net total swing prediction is shown in Figure 7.

(a) measurement point I

(b) measurement point II

(c) measurement point III

(d) measurement point IV

(e) measurement point V

(f) measurement point VI

(g) measurement point VII

(h) measurement point VIII

Figure 7. Comparison of predictions by the model

It can be seen from Figure 6 that the prediction model of net total swing based on the improved GM(1, 1) axis proposed in this paper is closer to the actual net total swing of the measurement point than the predicted value of other algorithms. To accurately analyze the prediction accuracy of several models, the error comparison analysis of the improved GM(1, 1), GM(1, 1), BP, SVM, and RF is carried out. In addition, Due to the randomness of the neural network, each structural model has to go through ten times training and verification during parameter setting. The errors obtained are shown in table 2, and the comparison of the average errors of the eight measurement points is shown in Figure 8.

algorithm	Measurement Point I (%)	Measurement Point II (%)	Measurement point III (%)	Measurement Point IV (%)	Measurement Point V (%)	Measurement point VI (%)	Measurement point VII (%)	Measurement point VIII (%)	Average error (%)
Improved GM	11.27	18.3	17.04	22.48	25.23	20.99	12.62	17.25	18.1475
GM	49.69	176.45	111.69	56.13	25.35	107.17	42.14	53.26	77.735
BP	22.68	55.04	41.32	181.39	41.32	129.61	95.49	102.03	83.61
SVM	43.71	61.59	169.24	119.82	49.05	162.83	169.74	71.06	105.88
RF	17.17	54.02	40.16	85.93	22.92	87.68	144.26	86.87	67.37625

Table 2. Error analysis of data predicted by different models

Figure8. Comparison of average errors

As can be seen from Figure 8, compare with traditional GM(1, 1), BP, SVM, and RF, after weakening the axis net total swing with oscillations combined with accelerated translation and mean transformation, the average error of GM(1, 1) model established by sum factor transformation is smaller than that of other models, it is indicating high prediction accuracy. The problem of the large prediction error of other algorithms for oscillating wave sequence is solved.

The sine function curve of axes net total swing is constructed by the improved GM(1, 1) eight-point net total swing prediction. The axial net total swing sine functions for the 71st-80th are shown in table 3.

number of times	the net total swing sine function
71	-0.03513 + 22.09*cos(x*0.01756)+40.96*sin(x*0.01756)
72	0.3865+14.04*cos(x*0.01716)-21.26*sin(x*0.01716)
73	0.293+12.47*cos(x*0.01721)-20.37*sin(x*0.01721)
74	0.02027 + 14.22*cos(x*0.01739) +20.43*sin(x*0.01739)
75	0.1227 + 18.76*cos(x*0.01727) + 13.77*sin(x*0.01727)
76	-0.2298 + 14.38*cos(x*0.01792)+15.78*sin(x*0.01792)
77	-0.1641 + 15.83*cos(x*0.01774) + 14.59*sin(x*0.01774)
78	-0.1342+14.02*cos(x*0.01756)-23.46*sin(x*0.01756)
79	0.2045 - 17.19*cos(x*0.01774) - 10.7*sin(x*0.01774)
80	0.09773 + 5.457*cos(x*0.01716) - 1.623*sin(x*0.01716)

Table 3. Sine function of net full swing for 71-80 sub-axes

The 71-80 times net total swing sine function is used as input of each method to search for the maximum value of net total swing sine function, and the maximum value obtained is the orientation adjustment. According to the results of 10 times of prediction, the obtained axis adjustment azimuth search results are shown in table 4.

number	Actual orientation (function maximum) ($^{\circ}$)	Traditional handwork ($^{\circ}$)	Error ($^{\circ}$)	Particle group (math) ($^{\circ}$)	Error ($^{\circ}$)	Simulated annealing ($^{\circ}$)	Error ($^{\circ}$)	Flamingos ($^{\circ}$)	Error ($^{\circ}$)
71	60	45	15	61.287	1.287	61.28	1.28	60.69	0.69
72	321	315	6	308.626	12.37	310.82	10.18	319.62	1.38
73	289	315	26	305.735	16.73	308.1	19.1	288.45	0.55
74	59	45	14	55.361	3.639	55.115	3.885	57.46	1.54
75	40	45	5	36.664	3.336	37.665	2.335	38.29	1.71
76	48	45	3	46.416	1.584	46.324	1.676	46.56	1.44
77	67	90	23	41.976	25.02	49.624	17.37	65.44	1.56
78	294	315	21	299.034	5.034	296.08	2.08	293.33	0.67
79	223	215	8	208.475	14.52	210.69	12.31	221.48	1.52
80	340	315	25	349.825	9.825	349.03	9.03	340.69	0.69

Table 4. Search results for axial adjustment orientation

As can be seen from table 4, the search results of the simulated annealing algorithm are significantly smaller in error than the traditional manual and particle swarm methods, with high search accuracy and good stability. However, the search results that use the flamingo algorithm are close to the actual orientation and it has higher stability. In addition, the minimum error of the flamingo algorithm can reach 0.55, and the search results of each method compared with the actual orientation are shown in Figure 9.

Figure 9. Comparison of search results by method

In order to accurately analyze the accuracy of the axial adjustment orientation search of the centralized method, an error

comparison is performed as shown in table 5, as shown in Figure 10. From the table, it can be seen that from the selected 10 experimental results, the maximum error of traditional manual is 26, the minimum error is 3, and the average error is 14.6, the maximum error of particle swarm is 25.024, the minimum error is 1.287, and the average error is 9.33, the maximum error of simulated annealing is 19.1, the minimum error is 1.28, and the average error is 7.92, and the maximum error of flamingo is the maximum error of flamingo is 1.71, the minimum error is 0.55, and the average error is 1.26. After the flamingo search algorithm seeks better than traditional manual by 13.4, better than particle swarm by 8.07, and better than simulated annealing by 6.66. The flamingo search algorithm is more significant in the Net total swing of axis seeking.

Algorithm	Traditional		Particle group (math.)		Simulated annealing/mm		Flamingos/mm	
	handwork/mm		/mm					
the quantity used as a parameter	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN
numerical value	26	3	25.024	1.287	19.1	1.28	1.71	0.55
average error	14.6		9.33		7.92		1.26	

Table 5. Average error comparison table

Figure 10. Comparison of the average error of each method

Conclusion

Axis adjustment method based on machine learning is an effective axis adjustment method. In this paper, our goal is used machine learning algorithm to improve the overall axis adjustment efficiency and axis adjustment accuracy. The main work is as follows: I) A prediction model based on GM(1, 1) is proposed. The model includes accelerated translation and mean value transformation to pre-process the oscillatory axis net total swing series, so as to weaken the fluctuation of the axis net total swing series. II) Determine the orientation of axis adjustment, in this paper, an optimization model of axis adjustment orientation based on flamingo search algorithm is proposed, the experimental results show that the average error of the flamingo search algorithm is 1.26, which can save more than 80% of the time compared with manual axis orientation search. It shows that the flamingo search algorithm-based axis orientation adjustment proposed in this paper is practical. The original traditional labor consumption affects the power generation in an average of 5 days. According to the average power generation of 62,000 kW units: $62,000 \times 24 \text{ hours} \times 5 \text{ days} = 7.44 \text{ million kWh/unit}$, the method proposed in this paper can be shortened to 1 day, and it impact on power generation is only fifth of the previous one, that is, 1.488 million kWh/unit. Compared with the traditional manual stop axis maintenance operation, the axis measurement time in this paper can be reduced by more than 80%, the labor costs of each unit is calculated from the traditional 16 people, 5 days construction period, and 550 yuan per person: $5 \text{ days} \times 16 \text{ people} \times 550 \text{ yuan/day} = 44,000 \text{ yuan/unit}$, which is reduced to 11,100 yuan/unit for 2 people per day. It provides a reliable basis for finding the axis deviation and twists of the unit in time and finding solutions, it improves directly and indirectly, economic benefits, and it has wide application prospects and it is suitable for axis adjustment. It provides a strong guarantee for the safe and stable operation of the power grid and a solid technical support for the maintenance of hydro-generator unit.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due at the request of the hydropower company, the data used in this paper can not be made public because of the operation data of the hydropower company's own power equipment but are available from the corresponding author on reasonable request.

References

1. Chang E C, Wu R C, Zhu K, et al. Adaptive neuro-fuzzy inference system-based grey time-varying sliding mode control for power conditioning applications[J]. *Neural Computing & Applications*, 2016(3).
2. Xia X, Ni W. A novel failure analysis and diagnosis method for hydraulic-turbine generator unit[J]. *Journal of Vibroengineering*, 2016, 18(6):3568-3580.
3. Sheikhan M, Mohammadi N. Erratum to: Time series prediction using PSO-optimized neural network and hybrid feature selection algorithm for IEEE load data[J]. *neural computing & applications*, 2013, 23(3-4):1195.
4. Shaik N B, Pedapati S R, Othman A R, et al. An intelligent model to predict the life condition of crude oil pipelines using artificial neural networks[J]. *Neural Computing and Applications*, 2021(3).
5. Zhu W, Zhou J, Xia X, et al. A novel KICA-PCA fault detection model for condition process of hydroelectric generating unit[J]. *Measurement*, 2014, 58:197-206.
6. Leng X, Xu S. Research on Intelligent Control of Synchronous Generator Excitation System Based on Computer Technology[J]. *Journal of Physics: Conference Series*, 2021, 1992(3):032125 (4pp).
7. Mo L I, Guo P, Zhang L, et al. Uncertain and multi-objective programming models for crop planting structure optimization[J]. *Frontiers of Agricultural Science and Engineering*, 2016.
8. Xiong L, Liu J, Song B, et al. Deep learning compound trend prediction model for hydraulic turbine time series[J]. *International Journal of Low-Carbon Technologies*, 2021(3):3.
9. Guo Z, Zhao X, Chen Y, et al. Short-term passenger flow forecast of urban rail transit based on GPR and KRR[J]. *Intelligent Transport Systems, IET*, 2019.
10. Zhang X, Zhou J, Guo J, et al. Vibrant fault diagnosis for hydroelectric generator units with a new combination of rough sets and support vector machine[J]. *Expert Systems with Applications*, 2012, 39(3):2621-2628.
11. Fu W, Wang K, Li C, et al. Vibration trend measurement for hydropower generator based on optimal variational mode decomposition and LSSVM improved with chaotic sine cosine algorithm optimization[J]. *Measurement Science and Technology*, 2018.
12. Huang H, Qin A, Mao H, et al. The Prediction Method on the Early Failure of Hydropower Units Based on Gaussian Process Regression Driven by Monitoring Data[J]. *Applied Sciences*, 2020, 11(1):153.
13. Li G, Sun Y, He Y, et al. Short-Term Power Generation Energy Forecasting Model for Small Hydropower Stations Using GA-SVM[J]. *Mathematical Problems in Engineering*, 2014, 2014(2):1-9.
14. Lvarez-Troncoso R, Benetti C J, Sarr A B, et al. Impacts of hydroelectric power stations on Trichoptera assemblages in four rivers in NW Spain[J]. *Limnologica*, 2015, 53:35-41.
15. Cui Z, Wu J, Ding Z, et al. A hybrid rolling grey framework for short time series modelling[J]. *Neural Computing and Applications*, 2021.
16. Environmental I S. Outage Analysis and Maintenance Strategies in Hydroelectric Production[J]. *Encyclopedia of Information Science & Technology Third Edition*, 2015. Axis alignment orientation search
17. Gong X. Optimization of the Power Generation Control Process of Hydraulic Turbine Set Based on the Improved BFO-PSO Algorithm[J]. *Journal of Coastal Research*, 2019, 94(sp1):227.
18. Xiao, Zhihuai, He, et al. ACO-Initialized Wavelet Neural Network for Vibration Fault Diagnosis of Hydroturbine Generating Unit.[J]. *Mathematical Problems in Engineering*, 2015.
19. Cheng, YL, Wee, et al. Innovative reservoir sediments reuse and design for sustainability of the hydroelectric power plants[J]. *RENEW SUST ENERG REV*, 2014, 2014,36(-):212-219.
20. Chen Z, Yuan X, Tian H, et al. Improved gravitational search algorithm for parameter identification of water turbine regulation system[J]. *Energy Conversion & Management*, 2014, 78:306-315.
21. Thaeer H A, Bin S. Series division method based on PSO and FA to optimize Long-Term Hydro Generation Scheduling [J].

Sustainable Energy Technologies and Assessments, 2018, 29:106-118.

22. Secui D C, Hora C, Bendea G. Parameter estimation using a modified whale optimization algorithm for input-output curves of thermal and hydro power plants[J]. International Transactions on Electrical Energy Systems, 2020, 30(2):e12188.1-e12188.26.
23. Mortazavi A, Togan V, Moloodpoor M. Solution of structural and mathematical optimization problems using a new hybrid swarm intelligence optimization algorithm[J]. Advances in Engineering Software, 2019, 127(JAN.):106-123.
24. Dubey A K, Krishna R, Aravind S, et al. Prediction of Coronavirus Outbreak Based on Cuisines and Temperature Using Machine Learning Algorithms[J]. Social Science Electronic Publishing.
25. Niu W J, Feng Z K, Cheng C T, et al. A parallel multi-objective particle swarm optimization for cascade hydropower reservoir operation in southwest China[J]. Applied Soft Computing, 2018, 70:562-575.
26. Shang Y, Fan Q, Shang L, et al. Modified genetic algorithm with simulated annealing applied to optimal load dispatch of the Three Gorges Hydropower Plant in China[J]. Hydrological Sciences Journal/Journal des Sciences Hydrologiques, 2019, 64(9).
27. Cheng J, Wang L, Yan X. Cuckoo search algorithm with memory and the vibrant fault diagnosis for hydroelectric generating unit[J]. Engineering with Computers, 2018:1-16.
28. Cheng J, Zhu C, Fu W, et al. An Imitation medical diagnosis method of hydro-turbine generating unit based on Bayesian network[J]. Transactions of the Institute of Measurement and Control, 2019.
29. Cheng J, Wang L, Xiong Y. An improved cuckoo search algorithm and its application in vibration fault diagnosis for a hydroelectric generating unit[J]. Engineering Optimization, 2017:1-16.
30. Zhu, Wenlong, Xue, et al. A state tendency measurement for a hydro-turbine generating unit based on aggregated EEMD and SVR[J]. Measurement Science & Technology, 2015.
31. Pan, Luoping, Xueli. Characteristic parameter degradation prediction of hydropower unit based on radial basis function surface and empirical mode decomposition[J]. Journal of vibration and control: JVC, 2015, 21(11):2200-2211.
32. An X, Yang L, Pan L. Nonlinear prediction of condition parameter degradation trend for hydropower unit based on radial basis function interpolation and wavelet transform[J]. ARCHIVE Proceedings of the Institution of Mechanical Engineers Part C Journal of Mechanical Engineering Science 1989-1996 (vols 203-210), 2015, 229(18):1-9.
33. Wei, Dengfeng. network traffic prediction based on RBF neural network optimized by improved gravitation search algorithm[J]. Neural computing & applications, 2017, 28(8):2303-2312.
34. Tomasiello S, V Loia, Khaliq A. A granular recurrent neural network for multiple time series prediction[J]. Neural Computing and Applications, 2021(1):1-18.
35. Fu W, Wang K, Zhang C, et al. A hybrid approach for measuring the vibrational trend of hydroelectric units with enhanced multi-scale chaotic series analysis and optimized least squares support vector machine[J]. Transactions of the Institute of Measurement and Control, 2019, 41(15):4436-4449.
36. Chen D. Research on Traffic Flow Prediction in the Big Data Environment Based on the Improved RBF Neural Network[J]. IEEE Transactions on Industrial Informatics, 2017, PP(4):1-1.
37. Xia R, Huang J, Chen Y, et al. A study of the method of the thermal conductivity measurement for VIPs with improved RBF neural networks [J]. Measurement, 2016, 87:246-254.
38. Mia M, Dhar N R. Prediction and optimization by using SVR, RSM, and GA in hard turning of tempered AISI 1060 steel under effective cooling condition[J]. Neural Computing & Applications, 2017.
39. Y Sahraoui, Bendotti P, D'Ambrosio C. Real-world hydro-power unit-commitment: Dealing with numerical errors and feasibility issues[J]. Energy, 2019, 184:91-104.
40. Wang J, Guo M, Liu Y. Hydropower Unit Commitment with Nonlinearity Decoupled from Mixed Integer Nonlinear Problem[J]. Energy, 2018:S0360544218303621.
41. Daadaa M, S Séguin, Demeester K, et al. An optimization model to maximize energy generation in short-term hydropower unit commitment using efficiency points[J]. International Journal of Electrical Power & Energy Systems, 2021, 125:106419.
42. Vieira B, Viana A, Matos M, et al. A Multiple Criteria Utility-based Approach for the Unit Commitment with Wind Power and Pumped Storage Hydro[J]. Electric Power Systems Research, 2016, 131(FEB.):244-254.
43. Li Y, Li K, Yang Z, Yu Y, Xu R, Yang M. Stochastic optimal scheduling of demand response-enabled microgrids with renewable generations: An analytical-heuristic approach. Journal of Cleaner Production. 2022,330:129840.
44. Li Y, Han M, Yang Z, Li G. Coordinating flexible demand response and renewable uncertainties for scheduling of community

integrated energy systems with an electric vehicle charging station: a bi-level approach. IEEE Transactions on Sustainable Energy 2021, 12(4):2321-2331.

45. Li Y, Li J, Wang Y. Privacy-preserving spatiotemporal scenario generation of renewable energies: A federated deep generative learning approach. IEEE Transactions on Industrial Informatics, 2022, 18(4):2310-2320.

Acknowledgments

This work was supported by the Science and Technology Planning Projects of Jilin Province, No. 20210201134GX.

Author contributions

Author Contributions: Conceptualization, J.C. and Y.L.; methodology, Y.L.; software, D.-Y.C.; validation, Q.-Z.Y., and Z.-R.X.; formal analysis, Z.-R.X.; investigation, L.-Y.W., and D.-Y.C.; resources, L.-Y.W.; data curation, D.-Y.C. and Y.L.; writing—original draft, Y.L.; writing—review and editing, J.C.; visualization, Q.-Z.Y.; supervision, L.-Y.W.; project administration, D.-Y.C.

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J. C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.