J. Meas. Sci. Instrum., 2025, 16(2): 258-271. http://jmsi.publish.founderss.cn
DOI: 10.62756/jmsi.1674-8042.2025025 jmsi@nuc.edu.cn

Application of interval type-2 TSK FLS method based on IGWO algorithm in short-term photovoltaic power forecasting

LI Jun*, ZENG Yuxiang

School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

***Corresponding author:** LI Jun (lijun691201@mail.lzjtu.cn)

Received: May 20, 2024 **Revised:** July 15, 2024 **Accepted:** September 20, 2024

Abstract: For short-term PV power prediction, based on interval type-2 Takagi-Sugeno-Kang fuzzy logic systems (IT2 TSK FLS), combined with improved grey wolf optimizer (IGWO) algorithm, an IGWO-IT2 TSK FLS method was proposed. Compared with the type-1 TSK fuzzy logic system method, interval type-2 fuzzy sets could simultaneously model both intra-personal uncertainty and inter-personal uncertainty based on the training of the existing error back propagation (BP) algorithm, and the IGWO algorithm was used for training the model premise and consequent parameters to further improve the predictive performance of the model. By improving the gray wolf optimization algorithm, the early convergence judgment mechanism, nonlinear cosine adjustment strategy, and Levy flight strategy were introduced to improve the convergence speed of the algorithm and avoid the problem of falling into local optimum. The interval type-2 TSK FLS method based on the IGWO algorithm was applied to the real-world photovoltaic power time series forecasting instance. Under the same conditions, it was also compared with different IT2 TSK FLS methods, such as type I TSK FLS method, BP algorithm, genetic algorithm, differential evolution, particle swarm optimization, biogeography optimization, gray wolf optimization, etc. Experimental results showed that the proposed method based on IGWO algorithm outperformed other methods in performance, showing its effectiveness and application potential.

Key words: photovoltaic power; interval type-2 fuzzy logic system; grey wolf optimizer algorithm; forecast performance of model

0 Introduction

Solar energy has become an important component of renewable energy, but its intermittency and volatility have always hindered the development and utilization of solar energy[1]. However, due to the influence of meteorological factors such as solar radiation, environmental temperature, and air humidity, photovoltaic power generation has the disadvantages of randomness, intermittency, and uncontrollability, which will cause certain impacts on the power grid and bring difficulties and challenges to the optimization and scheduling of the power grid. On the other hand, the grid connection of large-scale photovoltaic power generation also brings challenges to the stable operation of the power grid and the reliability of electric energy. Therefore, reliable and accurate prediction of photovoltaic power generation has practical significance for improving the reliability of the power system[2,3].

Based on the difference in time scale, photovoltaic power forecasting can be divided into ultra-short-term, short-term, medium-term, and long-term forecasting. Among them, short-term photovoltaic power forecasting can provide a certain reference for the scheduling plan of the power system[4]. According to the different modeling methods, it can be divided into physical methods, statistical methods, and artificial intelligence methods[5].

Artificial intelligence methods are based on the historical power of the photovoltaic power station and meteorological data to establish a prediction model, without detailed physical parameters of the photovoltaic components. The model has strong anti-interference ability and high prediction accuracy, so it is the most widely used method. Typical methods include artificial neural networks[6], extreme learning machine[7], support vector machine[8], and other methods. On the other hand, the fuzzy logic system method characterized by IF-THEN rules is a powerful artificial intelligence method that has also been successfully applied to prediction modeling. A TSK FLS parameter based on gradient descent algorithm was designed and successfully applied to wind speed forecast[9]. Genetic algorithm (GA) and particle swarm optimization (PSO) were used to optimize the parameters of a type-1 of TSK

FLS, and it was successfully applied to sewage treatment system prediction[10].

Interval type-2 TSK fuzzy logic system (IT2 TSK FLS) was proposed[11]. Compared with the traditional type-1 TSK FLS, it can handle modeling uncertainty problems better[12,13]. As a powerful modeling tool, IT2 TSK FLS has been successfully applied to prediction modeling. The parameters of IT2 TSK FLS were optimized based on gradient descent algorithm, and it was successfully applied to the power prediction of photovoltaic power generation[14]. Similarly, intelligent optimization algorithms have also been successfully applied to parameter optimization of IT2 TSK FLS models. The genetic algorithm was applied to the parameter optimization design of IT2 TSK FLS, and it was successfully applied to the power load prediction, and the prediction effect was good[15].

On the other hand, grey wolf optimizer (GWO) algorithm was proposed, which simulated the ranking system and hunting behavior of a grey wolf pack to achieve multi-objective optimization[16]. It had advantages such as high optimization efficiency and few control parameters. The multi-layer perceptron was optimized based on GWO algorithm and tested on 5 classification datasets and 3 function approximation datasets[17]. Compared with GA, differential evolution (DE), and particle swarm optimization (PSO) algorithms, it has better optimization performance. The GWO algorithm was applied to the parameter design of type 1 TSK FLS controller, and it was successfully applied to the sun tracking system, and the control effect was good[18]. However, the traditional GWO algorithm suffers from issues such as linear convergence factor that cannot effectively balance the global exploration and local search processes, weak optimization ability, and easy early convergence to local extremes[19,20].

An improved grey wolf optimizer (IGWO) algorithm was proposed. By introducing an early convergence judgment mechanism and the Levy flight strategy, the algorithm could increase population diversity and escape local extreme values.

The GWO algorithm and the IGWO algorithm were applied to optimize the IT2 TSK FLS method for short-term photovoltaic power forecasting in a certain region. Under the same conditions, the proposed method was compared with the type-1 TSK FLS and IT2 TSK FLS to demonstrate its effectiveness.

1 Data pre-processing

1.1 Correlation analysis

It is necessary to explore the correlation between meteorological variables and photovoltaic output power, and select the meteorological variables that are most likely to cause changes in photovoltaic output power as inputs to the prediction model. Given two data sequences X and Y with length N, the Pearson correlation coefficient between X and Y, denoted by $\rho_{X,Y}$, can be expressed as

$$\rho_{X,Y}=\frac{\sum_{i=1}^{N}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{\sqrt{\sum_{i=1}^{N}\left(X_i-\bar{X}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)^2}}, \tag{1}$$

where X_i and Y_i denote the ith data in the sequence; $\bar{X}$ and $\bar{Y}$ denote the mean values of the sequences X and Y, respectively.

1.2 Fuzzy C-means algorithm

Compared with the hard clustering used in K-means, the FCM clustering algorithm incorporates fuzzy theory and realizes flexible fuzzy partitioning.

If the $\boldsymbol{v}_j\left(j=1,2,\cdots,c\right)$ is the jth cluster center, $u_j(\boldsymbol{x}_i)$ is the degree of the membership of the ith data in the jth class, which should satisfy

$$\sum_{j=1}^{c}u_j(\boldsymbol{x}_i)=1. \tag{2}$$

The objective function of clustering is defined as

$$J_{\mathrm{e}}=\sum_{j=1}^{c}\sum_{i=1}^{N}\left[u_j(\boldsymbol{x}_i)\right]^b\left\|\boldsymbol{x}_i-\boldsymbol{v}_j\right\|^2, \tag{3}$$

where b is a constant greater than 1, typically $b=2$.

Under the condition that Eq. (2) is satisfied, the minimum value of Eq. (3) is obtained. And the partial derivative of J_{e} with respect to $\boldsymbol{v}_j$ and $u_j(\boldsymbol{x}_i)$ is 0, we can get

$$\boldsymbol{v}_j=\frac{\sum_{i=1}^{N}\left[u_j(\boldsymbol{x}_i)\right]^b\boldsymbol{x}_i}{\sum_{i=1}^{N}\left[u_j(\boldsymbol{x}_i)\right]^b},\quad j=1,2,\cdots,c, \tag{4}$$

$$u_j(\boldsymbol{x}_i)=\frac{\left(1/\left\|\boldsymbol{x}_i-\boldsymbol{v}_j\right\|^2\right)^{1/(b-1)}}{\sum_{k=1}^{c}\left(1/\left\|\boldsymbol{x}_i-\boldsymbol{v}_k\right\|^2\right)^{1/(b-1)}},\quad j=1,2,\cdots,c. \tag{5}$$

Let $c_{\max}$ represent the maximum cluster in the FCM algorithm. Then, the clustering validity from 2 to $c_{\max}$ is evaluated using the Davies-Bouldin index (DBI). Smaller DBI values indicate better clustering performance, and the optimal number of clusters can be determined. The DBI is calculated by

$$DBI=\frac{1}{c}\sum_{i=1}^{c}\max_{j\neq i}\left(\frac{D_i+D_j}{d_{ij}}\right), \tag{6}$$

where D_i and D_j represent the average distance from the ith and jth data to the corresponding cluster center, d_{ij} represents the distance between the cluster centers of class i and cluster j.

2 Interval type-2 TSK FLS

The interval type-2 TSK FLS, abbreviated as IT2 TSK FLS, contains three types: A2-C1, A2-C0, and A1-C1, with A and C being abbreviations for the antecedents and consequents, respectively.

2.1 Interval A2-C1 type-2 TSK FLS

Assume that given N sets of input-output data pairs, i.e.,

$$\left(x^{(1)}:y^{(1)}\right),\cdots,\left(x^{(N)}:y^{(N)}\right)=\left\{x_1^{(t)},\cdots,x_p^{(t)}:y^{(t)}\right\}_{t=1}^{N}, \quad (7)$$

where $x_1\in X_1,\cdots,x_p\in X_p$ is the input to the model and $y\in Y$ is the output. An interval type-2 TSK FLS with M rules can be designed, and each of its rules has p antecedents, the IF-THEN of the ith rule R^i can be described as

$$R^i\text{: IF } x_1 \text{ is } \tilde{F}_1^i \text{ and } x_2 \text{ is } \tilde{F}_2^i\cdots x_p \text{ is } \tilde{F}_p^i,$$
$$\text{then } Y_i=C_0^i+C_1^ix_1+C_2^ix_2+\cdots+C_p^ix_p, \quad (8)$$

where $i=1,2,\cdots,M$, $C_j^i\left(j=0,1,\cdots,p\right)$ are type-1 fuzzy set; the output of the ith rule is a linear combination of type-1 fuzzy sets, also type-1 fuzzy sets; $\tilde{F}_k^i$ is the antecedent interval type-2 fuzzy set corresponding to $x^{(t)}$. Such rules take into account both the uncertainty of the membership function of the antecedent and the uncertainty of the parameters of the consequent. The $\tilde{F}_k^i$ in the rule is described by the kth Gaussian primary membership function with uncertain mean, i.e.,

$$\mu_{\tilde{F}_k^i}(x_k)=\exp\left\{-\frac{1}{2}\left(\frac{x_k-m_k}{\sigma_{\tilde{F}_k^i}}\right)^2\right\},\ m_{\tilde{F}_k^i}\in\left[m_{\tilde{F}_{k1}^i},m_{\tilde{F}_{k2}^i}\right], \quad (9)$$

where $k=1,\cdots,p$, $\mu_{\tilde{F}_k^i}(x_k)=\left[\underline{\mu}_{\tilde{F}_k^i}(x_k),\bar{\mu}_{\tilde{F}_k^i}(x_k)\right]$, $\bar{\mu}_{\tilde{F}_k^i}(x_k)$ and $\underline{\mu}_{\tilde{F}_k^i}(x_k)$ are uncertain mean of the upper and lower membership functions, respectively. $m_{\tilde{F}_{k1}^i}$, $m_{\tilde{F}_{k2}^i}$, $\sigma_{\tilde{F}_k^i}$ are abbreviated to $m_{k1}^i,m_{k2}^i,\sigma_k^i$, i.e.,

$$\bar{\mu}_{\tilde{F}_k^i}(x_k)=\begin{cases}N\left(m_{k1}^i,\sigma_k^i;x_k\right), & x_k<m_{k1}^i,\\ 1, & m_{k1}^i\leqslant x_k\leqslant m_{k2}^i,\\ N\left(m_{k2}^i,\sigma_k^i;x_k\right), & x_k>m_{k2}^i,\end{cases} \quad (10)$$

$$\underline{\mu}_{\tilde{F}_k^i}(x_k)=\begin{cases}N\left(m_{k1}^i,\sigma_k^i;x_k\right), & x_k\leqslant\dfrac{m_{k1}^i+m_{k2}^i}{2},\\ N\left(m_{k2}^i,\sigma_k^i;x_k\right), & x_k>\dfrac{m_{k1}^i+m_{k2}^i}{2},\end{cases} \quad (11)$$

where $N(m,\sigma;x_k)=\exp\left[-\frac{1}{2}\left(\frac{(x_k-m)}{\sigma}\right)^2\right]$.

$C_j^i=\left[c_j^i-s_j^i,c_j^i+s_j^i\right]$ in rule R^i, where c_j^i denotes the center of fuzzy set C_j^i, and s_j^i denotes the spread of fuzzy set C_j^i.

Theorem 1[21] In an IT2 TSK FLS with minimum t-norm, (a) the firing set $F^i(x)$ of the rule R^i is defined as

$$F^i(x)=\left[\underline{f}^i(x),\bar{f}^i(x)\right], \quad (12)$$

where $\underline{f}^i(x)$ and $\bar{f}^i(x)$ can be described as

$$\begin{cases}\underline{f}^i(x)=\underline{\mu}_{\tilde{F}_1^i}(x_1)*\cdots*\underline{\mu}_{\tilde{F}_p^i}(x_p),\\ \bar{f}^i(x)=\bar{\mu}_{\tilde{F}_1^i}(x_1)*\cdots*\bar{\mu}_{\tilde{F}_p^i}(x_p),\end{cases} \quad (13)$$

where $*$ denotes a t-norm such as minimum or product.

(b) The consequent part Y^i of the rule R^i is an interval type-1 set, $Y^i=\left[y_l^i,y_r^i\right]$, where

$$\begin{cases}y_l^i=\sum_{k=1}^{p}c_k^ix_k+c_0^i-\sum_{k=1}^{p}|x_k|s_k^i-s_0^i,\\ y_r^i=\sum_{k=1}^{p}c_k^ix_k+c_0^i+\sum_{k=1}^{p}|x_k|s_k^i+s_0^i.\end{cases} \quad (14)$$

Correspondingly, the output of the interval type-2 TSK FLS can be simplified as

$$Y_{\text{TSK},2}=[y_l,y_r]=$$
$$\int_{y^1\in[y_l^1,y_r^1]}\cdots\int_{y^M\in[y_l^M,y_r^M]}\int_{f^1\in[\underline{f}^1,\bar{f}^1]}\cdots\int_{f^M\in[\underline{f}^M,\bar{f}^M]}1\Big/\frac{\sum_{i=1}^{M}f^iy^i}{\sum_{i=1}^{M}f^i}, \quad (15)$$

where $\underline{f}^i$, $\bar{f}^i$, y_l, y_r are calculated by Eqs. (13) and (14). The corresponding output of defuzzification is

$$y_{\text{TSK},2}(x)=\frac{y_l+y_r}{2}. \quad (16)$$

2.2 Interval A2-C0 type-2 TSK FLS

If the consequent is considered to be a clear number, that is the special case, which is the interval type-2 TSK FLS A2-C0, then the rule R^i can be described as

$$R^i\text{: IF } x_1 \text{ is } F_1^i \text{ and } x_2 \text{ is } F_2^i\cdots x_p \text{ is } F_p^i,$$
$$\text{then } Y_i=c_0^i+c_1^ix_1+c_2^ix_2+\cdots+c_p^ix_p. \quad (17)$$

In **Theorem 1**, only (a) still applies. Eq. (15) is simplified to

$$Y_{\text{TSK},2}(x)=\left[y_l,y_r\right]=\int_{f^1\in\left[\underline{f}^1,\bar{f}^1\right]}\cdots\int_{f^M\in\left[\underline{f}^M,\bar{f}^M\right]}1\Big/\frac{\sum_{i=1}^{M}f^iy^i}{\sum_{i=1}^{M}f^i}. \quad (18)$$

Except for ${y_l}^i={y_r}^i=y^i$, the procedure to compute is the same as for the A2-C1 case.

2.3 Interval A1-C1 type-2 TSK FLS

If the consequent set is a type-1 fuzzy set, and the

antecedent is also a special case of a type-1 fuzzy set, that is the interval type-2 TSK FLS A1-C1, then the rule R^i can be described as

$$R^i\text{: IF } x_1 \text{ is } F_1^i \text{ and } x_2 \text{ is } F_2^i \cdots x_p \text{ is } F_p^i,$$
$$\text{then } Y_i = C_0^i + C_1^i x_1 + C_2^i x_2 + \cdots + C_p^i x_p. \quad (19)$$

At this time, **Theorem 1** still applies, the firing set of the ith rule is a clear number, and Eq. (15) can be simplified to

$$Y = \left[\frac{\sum_{i=1}^{M} f^i y_l^i}{\sum_{i=1}^{M} f^i}, \quad \frac{\sum_{i=1}^{M} f^i y_r^i}{\sum_{i=1}^{M} f^i} \right]. \quad (20)$$

The defuzzification output from Eqs. (15) and (20) is

$$y_{\mathrm{TSK},2}(x) = \frac{\sum_{i=1}^{M} f^i \left(\sum_{k=1}^{p} c_k^i x_k + c_0^i \right)}{\sum_{i=1}^{M} f^i}, \quad (21)$$

where c_k^i is the center of a type-1 fuzzy set C_k^i.

If we consider the special case where the consequent set is a crisp number and the premise set is a type-1 fuzzy set, i.e., A1-C0, the rule R^i can be described as

$$R^i\text{: IF } x_1 \text{ is } F_1^i \text{ and } x_2 \text{ is } F_2^i \cdots x_p \text{ is } F_p^i,$$
$$\text{then } Y_i = c_0^i + c_1^i x_1 + c_2^i x_2 + \cdots + c_p^i x_p. \quad (22)$$

The defuzzification output is

$$y_{\mathrm{TSK},1}(x) = \frac{\sum_{i=1}^{M} f^i (c_0^i + c_1^i x_1 + \cdots + c_p^i x_p)}{\sum_{i=1}^{M} f^i}. \quad (23)$$

2.4 Design of BP algorithm for interval type-2 TSK fuzzy logic systems

Given N sets of input-output data pairs, $(x_1^{(t)}, \cdots, x_p^{(t)}; y^{(t)})$, the goal is to design an IT2 TSK FLS and update its design parameters so that the error function in Eq. (24) is minimized within E training epochs.

$$e^{(t)} = \frac{1}{2} \left[y_{\mathrm{TSK},2}(\boldsymbol{x}^{(t)}) - y^{(t)} \right]^2. \quad (24)$$

The specific implementation steps of the backpropagation (BP) algorithm for IT2 TSK FLS method from A2-C1 are as follows.

Step 1 Given the rule number M, FCM algorithm is used to cluster the data, and all fuzzy pre and post parameters are initialized with random values.

Step 2 Set the counter of training epoch to 1, i.e., $count = 1$.

Step 3 Set the counter of the training data to unity, i.e., $t = 1$.

Step 4 Apply the input $x^{(t)} \in R^p$ to the IT2 TSK FLS and compute the total firing interval and consequent of each rule, i.e., $\underline{f}^i$ and $\overline{f}^i$ are computed by Eqs. (13) and (14).

Step 5 Use the four-step iteration method to calculate y_l and y_r, reorder and renumber the M rules, and establish L and R. Then y_l and y_r are respectively represented as

$$y_l = \frac{\sum_{i=1}^{M} f_l^i y_l^i}{\sum_{i=1}^{M} f_l^i} = \frac{\sum_{i=1}^{L} \overline{f}^i y_l^i + \sum_{j=L+1}^{M} \underline{f}^j y_l^j}{\sum_{i=1}^{L} \overline{f}^i + \sum_{j=L+1}^{M} \underline{f}^j} = y_l\left(\overline{f}^1, \cdots, \overline{f}^L, \underline{f}^{L+1}, \cdots, \underline{f}^M, y_l^1, \cdots, y_l^M \right), \quad (25)$$

$$y_r = \frac{\sum_{i=1}^{M} f_r^i y_r^i}{\sum_{i=1}^{M} f_r^i} = \frac{\sum_{i=1}^{R} \underline{f}^i y_l^r + \sum_{j=R+1}^{M} \overline{f}^j y_r^j}{\sum_{i=1}^{R} \underline{f}^i + \sum_{j=R+1}^{M} \overline{f}^j} = y_r\left(\underline{f}^1, \cdots, \underline{f}^R, \overline{f}^{R+1}, \cdots, \overline{f}^M, y_l^1, \cdots, y_l^M \right). \quad (26)$$

Step 6 The defuzzied output of IT2 TSK FLS is calculated as

$$y_{\mathrm{TSK},2}(x^{(t)}) = \frac{y_l(x^{(t)}) + y_r(x^{(t)})}{2}. \quad (27)$$

Step 7 Determine the display dependency of the membership functions for y_l and y_r. First, as shown in Eq. (13), $\underline{f}^i$ is determined by $\underline{\mu}_{\bar{F}_1^i}(x_1), \cdots, \underline{\mu}_{\bar{F}_p^i}(x_p)$, and $\overline{f}^i$ is determined by $\overline{\mu}_{\bar{F}_1^i}(x_1), \cdots, \overline{\mu}_{\bar{F}_p^i}(x_p)$. Additionally, y_l^i and y_r^i are determined by c_j^i and s_j^i, respectively, for $j = 0, 1, \cdots, p$. Therefore, based on Eqs. (25) and (26), we can obtain

$$y_l = y_l\Big[\overline{\mu}_{\bar{F}_1^1}(x_1), \cdots, \overline{\mu}_{\bar{F}_p^1}(x_p), \cdots, \overline{\mu}_{\bar{F}_p^L}(x_1), \cdots, \overline{\mu}_{\bar{F}_p^L}(x_p), \underline{\mu}_{\bar{F}_1^{L+1}}(x_1), \cdots, \underline{\mu}_{\bar{F}_p^{L+1}}(x_p), \cdots, \underline{\mu}_{\bar{F}_1^M}(x_1), \cdots, \underline{\mu}_{\bar{F}_p^M}(x_p), c_0^1, \cdots, c_p^1, \cdots, c_0^M, \cdots, c_p^M, s_0^1, \cdots, s_p^1, \cdots, s_0^M, \cdots, s_p^M \Big], \quad (28)$$

$$y_r = y_r\Big[\underline{\mu}_{\bar{F}_1^1}(x_1), \cdots, \underline{\mu}_{\bar{F}_p^1}(x_p), \cdots, \underline{\mu}_{\bar{F}_1^R}(x_1), \cdots, \underline{\mu}_{\bar{F}_p^R}(x_p), \overline{\mu}_{\bar{F}_1^{R+1}}(x_1), \cdots, \overline{\mu}_{\bar{F}_p^{R+1}}(x_p), \cdots, \overline{\mu}_{\bar{F}_1^M}(x_1), \cdots, \overline{\mu}_{\bar{F}_p^M}(x_p), c_0^1, \cdots, c_p^1, \cdots, c_0^M, \cdots, c_p^M, s_0^1, \cdots, s_p^1, \cdots, s_0^M, \cdots, s_p^M \Big]. \quad (29)$$

Step 8 Based on the error function in Eq. (24), the Gaussian primary membership functions is considered with uncertain means shown in Eqs. (10) and (11), and gradient descent is used to adjust the parameters of the active branches of the antecedent membership functions and the parameters in the consequent.

Because the recursive updates of the upper and lower membership functions $m_{\bar{F}_{k1}^i}$ and $m_{\bar{F}_{k2}^i}$ are the same,

denoted as $m_{\bar{F}_k^i}$, the recursive formula for all design parameters in the A2-C1 type IT2 TSK FLS is

$$m_{\bar{F}_k^i}(t+1)=$$
$$m_{\bar{F}_k^i}(t)-\frac{1}{2}\alpha\left[y_{\mathrm{TSK},2}\left(x^{(t)}\right)-y^{(t)}\right]\cdot$$
$$\left[y_{\mathrm{l}}^{i}(t)-y_{\mathrm{TSK},2}\left(x^{(t)}\right)\right]\left[\frac{x_{\mathrm{k}}^{(t)}-m_{\bar{F}_k^i}(t)}{\sigma_{\bar{F}_k^i}^2(t)}\right]\phi_l\left(x^{(t)}\right), \quad (30)$$

$$\sigma_{\bar{F}_k^i}(t+1)=$$
$$\sigma_{\bar{F}_k^i}(t)-\frac{1}{2}\alpha\left[y_{\mathrm{TSK},2}\left(x^{(t)}\right)-y^{(t)}\right]\cdot$$
$$\left[y_{\mathrm{l}}^{i}(t)-y_{\mathrm{TSK},2}\left(x^{(t)}\right)\right]\left[\frac{\left(x_{\mathrm{k}}^{(t)}-m_{\bar{F}_k^i}(t)\right)^2}{\sigma_{\bar{F}_k^i}^3(t)}\right]\phi_l\left(x^{(t)}\right), \quad (31)$$

$$c_{j+1}(t+1)=c_{j+1}(t)-$$
$$\frac{1}{2}\alpha\left[y_{\mathrm{TSK},2}\left(x^{(t)}\right)-y^{(t)}\right]x_j^{(t)}\left(\phi_l\left(x^{(t)}\right)+\phi_r\left(x^{(t)}\right)\right), \quad (32)$$

$$s_{j+1}(t+1)=s_{j+1}(t)+$$
$$\frac{1}{2}\alpha\left[y_{\mathrm{TSK},2}\left(x^{(t)}\right)-y^{(t)}\right]\left|x_j^{(t)}\right|\left(\phi_l\left(x^{(t)}\right)-\phi_r\left(x^{(t)}\right)\right), \quad (33)$$

where $j=0,1,\cdots,p-1$; $x_0^{(t)}=1$.

Step 9 Set $t=t+1$. If $t=N+1$, go to step 10, otherwise, go to step 4.

Step 10 Set $count=count+1$. If $count=E$, stop, otherwise, go to step 3.

Once the active branches of the antecedent membership function is determined, the partial derivatives of the upper and lower membership functions with respect to their parameters need to be calculated, and the partial derivatives of m_{k1}^i, for example, is

$$m_{k1}^i(t+1)=m_{k1}^i(t)-$$
$$\alpha\left[\frac{\partial e^{(t)}}{\partial y_{\mathrm{TSK},2}\left(x^{(t)}\right)}\frac{\partial y_{\mathrm{TSK},2}\left(x^{(t)}\right)}{\partial y_{\mathrm{l}}}\frac{\partial y_{\mathrm{l}}}{\partial m_{k1}^i}+\right.$$
$$\left.\frac{\partial e^{(t)}}{\partial y_{\mathrm{TSK},2}\left(x^{(t)}\right)}\frac{\partial y_{\mathrm{TSK},2}\left(x^{(t)}\right)}{\partial y_r}\frac{\partial y_r}{\partial m_{k1}^i}\right]\Bigg|_t, \quad (34)$$

where $\alpha>0$, $\left.\frac{\partial e^{(t)}}{\partial y_{\mathrm{TSK},2}\left(x^{(t)}\right)}\right|_t=y_{\mathrm{TSK},2}\left(x^{(t)}-y^{(t)}\right)$, $\left.\frac{\partial y_{\mathrm{TSK},2}}{\partial y_l}\right|_t=\frac{1}{2}$, and we have

$$\begin{cases}\left.\frac{\partial y_l}{\partial m_{k1}^i}\right|_t=\left[\frac{\partial y_l}{\partial \bar{\mu}_{\bar{F}_k^i}\left(x_k^{(t)}\right)}\frac{\partial \bar{\mu}_{\bar{F}_k^i}\left(x_k^{(t)}\right)}{\partial m_{k1}^i}\right]\Bigg|_t,\\ \left.\frac{\partial y_r}{\partial m_{k1}^i}\right|_t=\left[\frac{\partial y_r}{\partial \underline{\mu}_{\bar{F}_k^i}\left(x_k^{(t)}\right)}\frac{\partial \underline{\mu}_{\bar{F}_k^i}\left(x_k^{(t)}\right)}{\partial m_{k1}^i}\right]\Bigg|_t,\end{cases} \quad (35)$$

$$\frac{\partial \bar{\mu}_{\bar{F}_k^i}\left(x_k^{(t)}\right)}{\partial m_{k1}^i}=$$
$$\begin{cases}\left(x_k^{(t)}-m_{k1}^i\right)\partial N\left(m_{k1}^i,\sigma;x_k^{(t)}\right)/\sigma^2, & x_k^{(t)}<m_{k1}^i,\\ 0, & m_{k1}^i\leqslant x_k^{(t)}\leqslant m_{k2}^i,\\ 0, & x_k^{(t)}>m_{k2}^i,\end{cases} \quad (36)$$

where $N(m_{k1}^i,\sigma,x_k^{(t)})=\exp\left[-\frac{1}{2}\left(\left(x_k^{(t)}-m_{k1}^i\right)/\sigma\right)^2\right]$.

3 Design of interval type-2 TSK FLS based on IGWO algorithm

Due to the susceptibility of the BP algorithm to local minima, it can affect the predictive performance of the model to some extent. Therefore, as a derivative-free optimization algorithm, the GWO algorithm is very attractive for tuning parameters in an IT2 fuzzy system, which has the characteristics of a simple overall structure, fast convergence speed, and high search efficiency. However, the basic GWO algorithm often encounters the problems of premature convergence and local optima due to the rapid decrease in population diversity during the terminal phase of population evolution, which limits its further application in engineering optimization. An improved grey wolf optimization algorithm (IGWO) was constructed by introducing three improvement strategies: premature convergence judgment mechanism, Levy flight strategy, and nonlinear convergence factor adjustment strategy, which is used to optimize the parameters of the interval type-2 TSK FLS.

3.1 GWO

The GWO algorithm is a new type of swarm intelligence optimization algorithm that simulates the social structure and hunting behavior of grey wolves to solve optimization problems. In the algorithm, the population is divided into four levels based on their fitness values, with the top three grey wolves being labeled as α, β and δ in order, and the rest labeled as ω. α, β and δ lead the rest of the population to approach the prey location (global optimum) and complete the hunting process, as shown in Fig.1.

The mathematical model for grey wolf hunting is

$$\boldsymbol{D}=\boldsymbol{C}\cdot\boldsymbol{X}_{\mathrm{P}}(T)-\boldsymbol{X}(T), \quad (37)$$

$$\boldsymbol{X}(T+1)=\boldsymbol{X}_{\mathrm{P}}(T)-\boldsymbol{A}\cdot\boldsymbol{D}, \quad (38)$$

where T is the current iteration number; $\boldsymbol{A}$ and $\boldsymbol{C}$ are coefficient vectors; $\boldsymbol{X}_{\mathrm{P}}$ is the prey position vector, i.e., the global optimal solution; $\boldsymbol{X}(T)$ is the position vector of the gray wolf, i.e., the vector to be optimized; $\boldsymbol{D}$ is the

distance vector between the gray wolf and the prey.

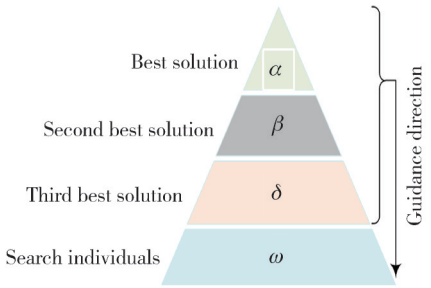


Fig. 1 Hierarchy of grey wolves

The coefficient vectors $\boldsymbol{A}$ and $\boldsymbol{C}$ are

$$\boldsymbol{A}=2\boldsymbol{a}\cdot\boldsymbol{r}_1-\boldsymbol{a}, \tag{39}$$

$$\boldsymbol{C}=2\boldsymbol{r}_2, \tag{40}$$

$$\boldsymbol{a}=\boldsymbol{a}_0\left(1-T/T_0\right), \tag{41}$$

where T_0 is the maximum number of iterations; $\boldsymbol{a}$ is the convergence factor, and $\boldsymbol{a}_0$ is the initial convergence factor; $\boldsymbol{r}_1,\boldsymbol{r}_2$ are a random vector within the range of $[0,1]$.

Gray wolf hunting location update formulas are

$$\begin{cases}\boldsymbol{D}_\alpha=\left|\boldsymbol{C}_1\cdot\boldsymbol{X}_\alpha-\boldsymbol{X}\right|,\\ \boldsymbol{D}_\beta=\left|\boldsymbol{C}_2\cdot\boldsymbol{X}_\beta-\boldsymbol{X}\right|,\\ \boldsymbol{D}_\delta=\left|\boldsymbol{C}_3\cdot\boldsymbol{X}_\delta-\boldsymbol{X}\right|,\end{cases} \tag{42}$$

$$\begin{cases}\boldsymbol{X}_1=\boldsymbol{X}_\alpha-\boldsymbol{A}_1\cdot(\boldsymbol{D}_\alpha),\\ \boldsymbol{X}_2=\boldsymbol{X}_\beta-\boldsymbol{A}_2\cdot(\boldsymbol{D}_\beta),\\ \boldsymbol{X}_3=\boldsymbol{X}_\delta-\boldsymbol{A}_3\cdot(\boldsymbol{D}_\delta),\end{cases} \tag{43}$$

$$\boldsymbol{X}(T+1)=\frac{\boldsymbol{X}_1+\boldsymbol{X}_2+\boldsymbol{X}_3}{3}, \tag{44}$$

where $\boldsymbol{X}_\alpha,\boldsymbol{X}_\beta,\boldsymbol{X}_\delta$ represent the position vectors of α,β,δ; $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ represent the directional vectors of ω towards α,β,δ in the gray wolf population; $\boldsymbol{X}(T+1)$ represents the final position, which is the direction towards the prey that the remaining wolves ω move towards.

3.2 IGWO

3.2.1 Early maturity convergence detection mechanism

The GWO algorithm still tends to fall into premature convergence, i. e., the phenomenon of premature convergence. To address this issue, an early convergence detection mechanism is introduced, and the population fitness variance (σ^2) is used as the evaluation index to measure the degree of aggregation of individual members in the population. This index is defined as[22]

$$\sigma^2=\frac{1}{S}\sum_{i=1}^{S}\left(\frac{J_i-J_{\text{avg}}}{J}\right)^2, \tag{45}$$

where S represents the population size; J_i represents the fitness value of the ith gray wolf; J_{avg} represents the average fitness value of the population; J is the normalization factor, which limits the size of σ^2. They are defined as

$$J_{\text{avg}}=\frac{1}{S}\sum_{i=1}^{S}J_i, \tag{46}$$

$$J=\max\left\{1,\max\left|J_i-J_{\text{avg}}\right|\right\}. \tag{47}$$

Eq. (45) indicates that the smaller the convergence factor σ^2, the higher the degree of aggregation among individuals in the population. This aggregation will lead to a lack of diversity in the population, causing the algorithm to fall into a local optimum. To avoid this premature convergence phenomenon, early maturity processing is performed when $\sigma^2<c_0$ (c_0 is a user-defined constant).

3.2.2 Levy flight strategy

To address the problem of the algorithm getting stuck in local optima, the Levy flight strategy is introduced to increase the diversity of the grey wolf population. If the premature convergence condition ($\sigma^2<c_0$) is met, the Levy flight strategy is used to fine-tune the positions of the individuals in the population. The random step size of Levy flight can be represented as[23]

$$Levy(\beta)=\frac{\mu}{|v|^{\frac{1}{\beta}}},\ (\beta=1.5), \tag{48}$$

where $Levy(\beta)$ is the step size of Levy flight; μ,v obey the random number of normal distribution, i.e.,

$$\mu\sim N(0,\sigma_\mu{}^2),\quad v\sim N(0,\sigma_v{}^2), \tag{49}$$

$$\text{and}\begin{cases}\sigma_\mu=\left\{\dfrac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma\left[(1+\beta)/2\right]\beta\times 2^{(\beta-1)/2}}\right\}^{1/\beta},\\ \sigma_v=1.\end{cases} \tag{50}$$

When the algorithm falls into premature convergence ($\sigma^2<c_0$), the Levy flight strategy is introduced to improve the position update formula for grey wolf optimization.

$$\begin{cases}\boldsymbol{X}_1'=\boldsymbol{X}_\alpha+Levy(\beta)\oplus\boldsymbol{\alpha}_1,\\ \boldsymbol{X}_2'=\boldsymbol{X}_\beta+Levy(\beta)\oplus\boldsymbol{\alpha}_2,\\ \boldsymbol{X}_3'=\boldsymbol{X}_\delta+Levy(\beta)\oplus\boldsymbol{\alpha}_3,\end{cases} \tag{51}$$

$$\begin{cases}\boldsymbol{\alpha}_1=0.01(\boldsymbol{X}_\alpha-\boldsymbol{X}),\\ \boldsymbol{\alpha}_2=0.01(\boldsymbol{X}_\beta-\boldsymbol{X}),\\ \boldsymbol{\alpha}_3=0.01(\boldsymbol{X}_\delta-\boldsymbol{X}),\end{cases} \tag{52}$$

$$\boldsymbol{X}_{\text{new}}(T+1)=\frac{\boldsymbol{X}_1'+\boldsymbol{X}_2'+\boldsymbol{X}_3'}{3}, \tag{53}$$

where $\oplus$ is the term-by-term product; α is the step size information, which is used to control the range of Levy flight random search.

3.2.3 Nonlinear convergence factor adjustment strategy

For the basic GWO algorithm, the convergence factor $\boldsymbol{a}$ linearly decreases with increasing iterations, which hinders the algorithm's ability for global optimization. Hence, a nonlinear cosine convergence factor is introduced, i.e.,

$$\boldsymbol{a}=\frac{\boldsymbol{a}_0}{2}\left(\cos\left(\frac{\pi T}{T_0}\right)+1\right), \tag{54}$$

where the initial decrease rate of $\boldsymbol{a}$ is slow, which is beneficial for enhancing the algorithm's global exploration ability.

3.3 Building predictive models

To apply the IGWO algorithm to the parameters tuning of the IT2 TSK FLS, the position ($\boldsymbol{X}$) of a grey wolf individual is associated with the parameters to be optimized in the IT2 TSK FLS model. For the three different IT2 TSK FLS structures, A2-C1, A2-C0, and A1-C1, all the parameters to be optimized are represented by

$$\boldsymbol{X}_{(\mathrm{A2-C1})}=\Bigg\{\underbrace{\overbrace{\underline{m}_1^1,\bar{m}_1^1,\sigma_1^1}^{\text{Antecedent 1}}\cdots\overbrace{\underline{m}_p^1,\bar{m}_p^1,\sigma_p^1}^{\text{Antecedent }p}\overbrace{c_0^1,s_0^1,\cdots,c_p^1,s_p^1}^{\text{Consequent}}}_{\text{Rule 1}}\cdots$$
$$\underbrace{\overbrace{\underline{m}_1^M,\bar{m}_1^M,\sigma_1^M}^{\text{Antecedent 1}}\cdots\overbrace{\underline{m}_p^M,\bar{m}_p^M,\sigma_p^M}^{\text{Antecedent }p}\overbrace{c_0^M,s_0^M,\cdots,c_0^M,s_p^M}^{\text{Consequent}}}_{\text{Rule }M}\Bigg\}, \tag{55}$$

$$\boldsymbol{X}_{(\mathrm{A2-C0})}=\Bigg\{\underbrace{\overbrace{\underline{m}_1^1,\bar{m}_1^1,\sigma_1^1}^{\text{Antecedent 1}}\cdots\overbrace{\underline{m}_p^1,\bar{m}_p^1,\sigma_p^1}^{\text{Antecedent }p}\overbrace{c_0^1,\cdots,c_p^1}^{\text{Consequent}}}_{\text{Rule 1}}$$
$$\underbrace{\overbrace{\underline{m}_1^M,\bar{m}_1^M,\sigma_1^M}^{\text{Antecedent 1}}\cdots\overbrace{\underline{m}_p^M,\bar{m}_p^M,\sigma_p^M}^{\text{Antecedent }p}\overbrace{c_p^M,\cdots,c_p^M}^{\text{Consequent}}}_{\text{Rule }M}\Bigg\}, \tag{56}$$

$$\boldsymbol{X}_{(\mathrm{A1-C1})}=\Bigg\{\underbrace{\overbrace{m_1^1,\sigma_1^1}^{\text{Antecedent 1}}\cdots\overbrace{m_p^1,\sigma_p^1}^{\text{Antecedent }p}\overbrace{c_0^1,s_0^1,\cdots,c_p^1,s_p^1}^{\text{Consequent}}}_{\text{Rule 1}}\cdots$$
$$\underbrace{\overbrace{m_1^M,\sigma_1^M}^{\text{Antecedent 1}}\cdots\overbrace{m_p^M,\sigma_p^M}^{\text{Antecedent }p}\overbrace{c_0^M,s_0^M,\cdots,c_p^M,s_p^M}^{\text{Consequent}}}_{\text{Rule }M}\Bigg\}. \tag{57}$$

In the implementation process of the TSK FLS method optimized by IGWO algorithm, the fitness function to be maximized is

$$J(\boldsymbol{X})=\frac{1}{1+E_{\mathrm{RMSE}}}, \tag{58}$$

$$E_{\mathrm{RMSE}}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[y(i)-y_d(i)\right]^2}, \tag{59}$$

where E_{RMSE} is the root mean square error; $y(i)$ is the predicted output of the model; $y_d(i)$ is the actual value; N is the number of training set.

In summary, the implementation steps for optimizing the A2-C1 type TSK FLS method based on the IGWO algorithm are as follows.

Step 1 Import the training dataset, perform correlation analysis and clustering on the dataset, and divide it into several sub-datasets.

Step 2 Initialize parameters by setting the number of fuzzy rules M, initial values of antecedent and consequent parameters for the IT2 TSK FLS, maximum number of iterations T_0, and initializing coefficient vectors $\boldsymbol{a}$, $\boldsymbol{A}$, and $\boldsymbol{C}$. Initialize the grey wolf population $\boldsymbol{X}_j(j=1,2,\cdots,\mathrm{S})$ within the parameter search range.

Step 3 Calculate the fitness $J(\boldsymbol{X})$ of each grey wolf individual in the initial population using Eq. (58), sort in descending order, and save the best, second-best, and third-best individual solutions as $\boldsymbol{X}_\alpha$, $\boldsymbol{X}_\beta$, and $\boldsymbol{X}_\delta$.

Step 4 Calculate the non-linear convergence factor $\boldsymbol{a}$ using Eq. (54) and update $\boldsymbol{A}$ and $\boldsymbol{C}$ using Eqs. (39) and (40).

Step 5 Calculate the fitness variance σ^2 of the population using Eqs. (45) to (47), and use the early convergence judgment mechanism to determine if the algorithm has reached a local extreme value. If $\sigma^2<c_0$, proceed to Step 6, otherwise, jump to step 7.

Step 6 Build the Levy flight model. Calculate $Levy(\beta)$ using Eq. (48) and update the grey wolf population individual information $\boldsymbol{X}_{\mathrm{new}}$ using Eqs. (51) to (53).

Step 7 Update the individual information $\boldsymbol{X}$ of the grey wolf population using Eqs. (43) and (44).

Step 8 Update the fitness $J(\boldsymbol{X})$ of all individuals in the grey wolf population using Eq. (58), and update $\boldsymbol{X}_\alpha$, $\boldsymbol{X}_\beta$, and $\boldsymbol{X}_\delta$ accordingly.

Step 9 Set $T=T+1$, if $T<T_0$, jump to step 3, otherwise terminate the algorithm, the optimal solution $\boldsymbol{X}_\alpha$ is obtained, and convert it to the parameters of the A2-C1 type TSK FLS model using Eq. (55).

3.4 Evaluation metrics for predictive models

Normalization preprocessing is required for the training data, and the performance metrics of the model are measured using root mean square error E_{RMSE}, mean absolute error E_{MAE}, and coefficient of determination R^2.

$$E_{\mathrm{MAE}}=\frac{1}{N}\sum_{i=1}^{N}\left|y(i)-y_d(i)\right|, \tag{60}$$

$$R^2=1-\frac{\sum_{i=1}^{N}\left(y_d(i)-y(i)\right)^2}{\sum_{i=1}^{N}\left(y_d(i)-\bar{y}\right)^2}, \tag{61}$$

where $\bar{y}$ is the mean value of the training data set.

4 Application examples

The actual operating data and meteorological data of a photovoltaic power station in Alice Springs, Australia, from March 2017 to February 2019, were selected as the test object[24]. This dataset included the output power data (PV) for the whole year as well as 9 meteorological data: wind speed (WS), temperature (t), global horizontal irradiance (GHI), weather relative humidity (WRH), wind direction (WD), diffuse horizontal irradiance (DHI), daily rainfall (DR), global tilted irradiance (IGT), and diffuse tilted irradiance (IDT). The data were sampled at intervals of 1 hour.

4.1 Feature selection and training data set partitioning

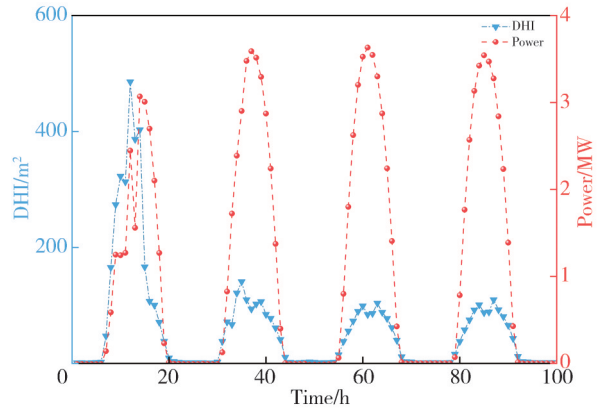
To identify the meteorological variables that have a significant impact on the output power of the photovoltaic station, a correlation analysis is needed between the power and meteorological data over a certain period of time. Table 1 shows the correlation measures between the photovoltaic output power and the meteorological variables at the Alice Springs station, as well as the statistical information of each meteorological variable and the photovoltaic output power. Based on the Pearson correlation coefficients presented in Table 1, the top seven meteorological variables with the strongest correlation are selected as inputs for the prediction model.

Table 1 Correlation analysis between photovoltaic power and meteorological factors

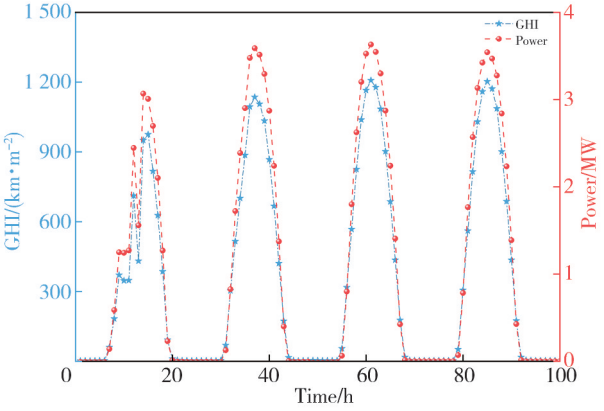
Meteorological variable	Correlation coefficient	Statistic information	
		Range	Average
WS/(m•s^{-1})	0.476 2	[0.127 5, 10.215 8]	1.967 2
t/(°C)	0.493 8	[−3.240 7, 43.138 1]	21.927 9
WRH/%	−0.440 9	[4.324 0, 101.703 7]	34.688 1
GHI/(W•m^{-2})	0.993 9	[1.312 0, 1226.890 9]	267.531
DHI/(W•m^{-2})	0.580 1	[0.418 8, 624.641 8]	51.758 4
WD/(°)	−0.068 3	[41.383 5, 1848.711]	283.294 0
DR/mm	−0.022 4	[0, 36.400 0]	0.376 7
IGT/(W•m^{-2})	0.993 5	[0.341 2, 1223.235 4]	291.199 1
IDT/(W•m^{-2})	0.608 3	[0.341 2, 1223.235 4]	53.236 5
PV/kW	—	[−0.016 9, 3.846 3]	0.905 4

For a more visual observation of the correlation between the selected meteorological variables and the PV output power, the curves containing both meteorological variables and PV output power are presented in Fig.2.

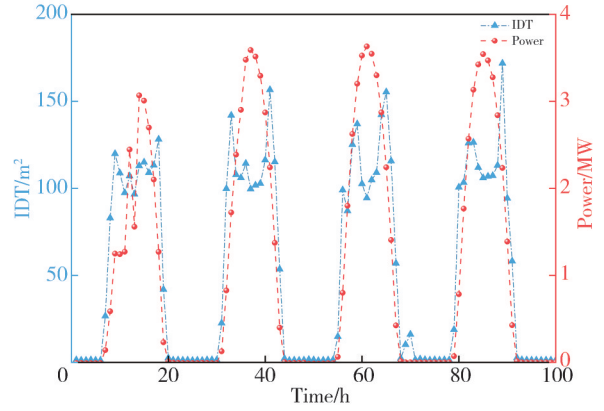
It can also be concluded that different meteorological variables have different ranges and metrics. Therefore, all data features are normalized to eliminate the influence of the magnitudes of the different features.



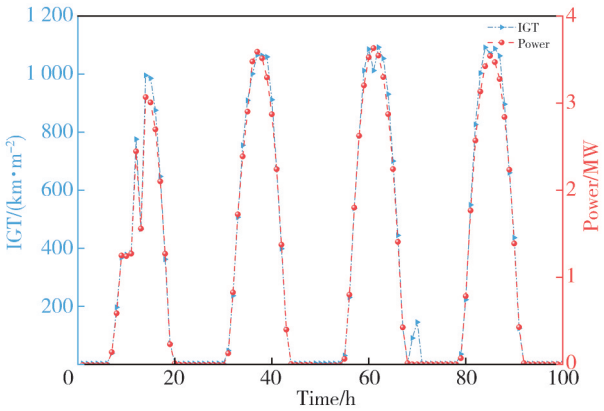
(a) Comparison between PV power and DHI



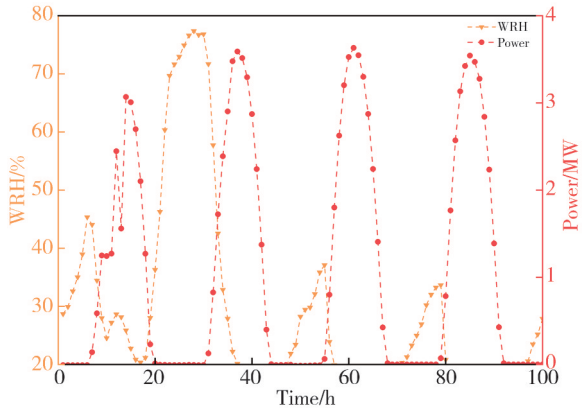
(b) Comparison between PV power and GHI



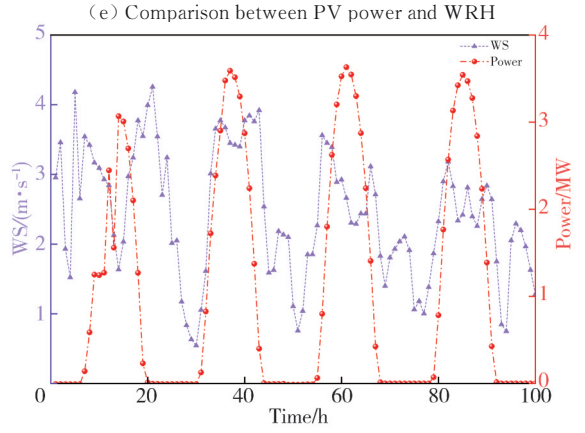
(c) Comparison between PV power and IDT



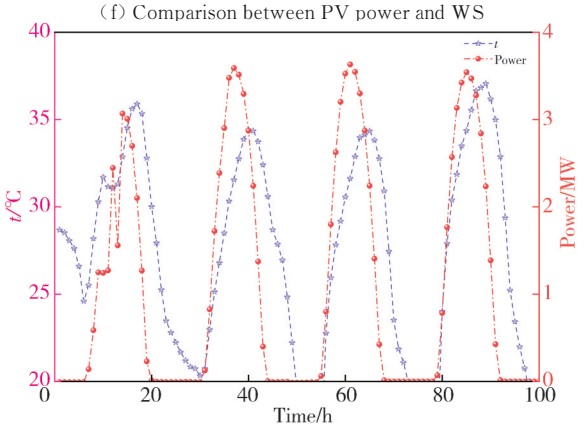
(d) Comparison between PV power and IGT



(e) Comparison between PV power and WRH



(f) Comparison between PV power and WS



(g) Comparison between PV power and t

Fig. 2 Comparison graphs between PV power and selected meteorological factors

Fig. 3 further presents the hourly average PV power values for each season. It can be seen that the trend of PV power with seasonal variation is very obvious. Based on this situation and the local climate conditions, the test dataset is divided into four datasets for each season: autumn (March to May), winter (June to August), spring (September to November), and summer (January, February, and December), in order to verify the predictive performance of the model under different seasons.

The preprocessed dataset is clustered into similar days using the FCM algorithm. To determine a reasonable number of clusters, considering the size of the dataset, c_{max} is set to 10, and c is allowed to range from 2 to 10. The corresponding values of the DBI are calculated using Eq. (6) and are shown in Fig.4. It can be seen that DBI = 0.527 9 reaches its minimum value when $c = 3$, indicating that setting the number of clusters to 3 is the most appropriate.

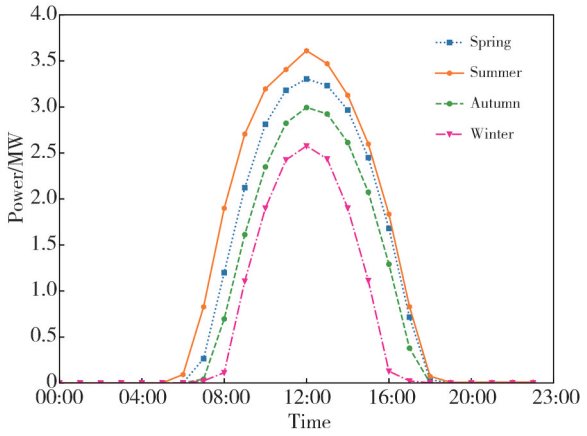


Fig. 3 Hourly mean photovoltaic power generation of photovoltaic plant in different seasons

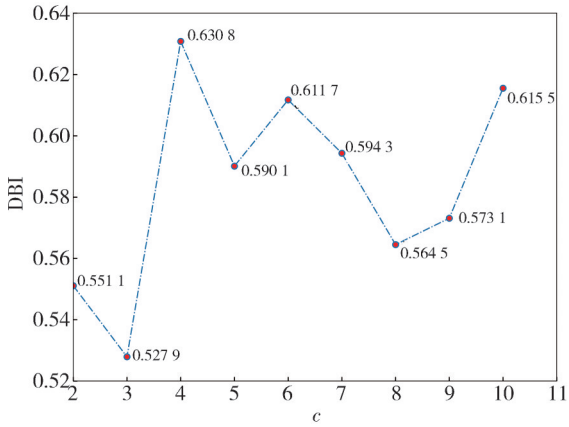
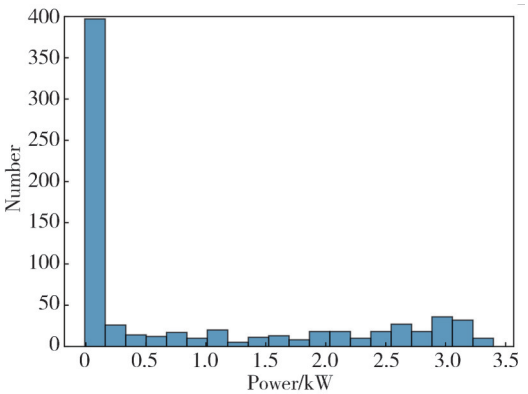
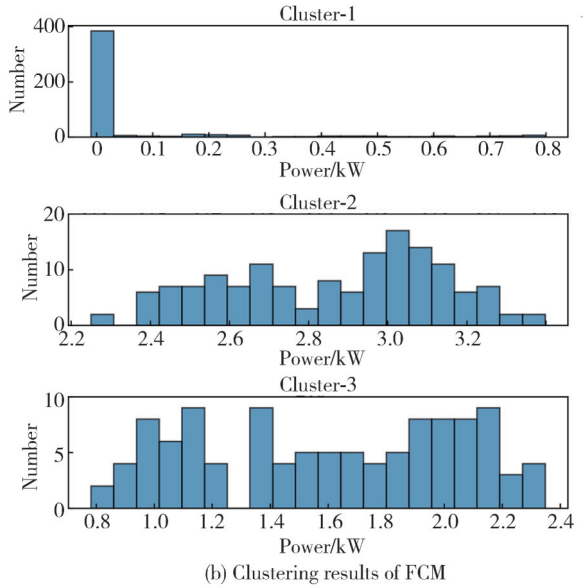


Fig. 4 Evaluation results by DBI for different cluster results

To display the clustering effect more intuitively, the power distribution histogram of the training dataset before clustering is shown in Fig.5 (a) using the April 2017 dataset as an example.



(a) Training dataset



(b) Clustering results of FCM

Fig. 5 Histogram of power distribution of training data set

The power distribution histograms of the three training data subsets after clustering are shown in Fig.5(b). It can be visually seen that the original training dataset has been successfully partitioned into different data partitions.

4.2 Simulation results

Different independent prediction models were established based on the final data partitions of each training dataset in different seasons. The models include SVM method and fuzzy logic methods based on BP, GA, DE, PSO, BBO, GWO, and IGWO algorithms for optimizing different TSK FLS models.

All experiments were conducted using Matlab R2020b on a PC equipped with an Intel(R) Core (TM) i7-12700H, 2.30 GHz processor for simulation computations.

For the sake of comparison, the fuzzy rule number M in different TSK FLS methods is set to 5. For the four fuzzy system models A1-C0, A1-C1, A2-C0, and A2-C1, the number of parameters to be optimized in the antecedent and consequent are shown in Table 2, and the initial values of the antecedent and consequent parameters for different methods are shown in Table 3.

Table 2 Parameters of different TSK FLS and number of parameters

Model	Parameter number	Designed parameters
A1-C0	$(3p+1)M$	m_k^i, σ_k^i, c_j^i
A1-C1	$(4p+2)M$	$m_k^i, \sigma_k^i, c_j^i, s_j^i$
A2-C0	$(4p+1)M$	$m_{k1}^i, m_{k2}^i, \sigma_k^i, c_j^i$
A2-C1	$(5p+2)M$	$m_{k1}^i, m_{k2}^i, \sigma_k^i, c_j^i, s_j^i$

Meanwhile, the hyperparameter settings for GA, DE, PSO, BBO, GWO, and IGWO algorithms are as follows. The population size for all algorithms is set to 60, and the maximum number of iterations is 40. Among them, the crossover probability for GA algorithm is set to 0.4 and the mutation probability is set to 0.01. The scaling factor for DE algorithm is set to 0.4 and the crossover probability is set to 0.6. The local and global learning factors for PSO algorithm are both set to 0.6, and the inertia factor is set to 0.3. For BBO algorithm, the initial mutation probability is set to 0.02, the maximum and minimum migration rates are both set to 1, and the elite retention number is set to 3. For GWO algorithm, the initial convergence factor $\boldsymbol{a}_0 = 2$ is set. For IGWO algorithm, the initial convergence factor $\boldsymbol{a}_0 = 2$ and the premature judgment constant $c_0 = 0.06$ are set.

Table 3 Initialization of model parameters

Model	Antecedent parameter	Consequent parameter
A1-C0	$m\in(-1,1),\sigma\in(2,4)$	$c\in(-1,1)$
A1-C1	$m\in(-1,1),\sigma\in(2,4)$	$c\in(-1,1),s\in(0,1)$
A2-C0	$\underline{m}\in(-1,0),\bar{m}\in(0,1),\sigma\in(2,4)$	$c\in(-1,1)$
A2-C1	$\underline{m}\in(-1,0),\bar{m}\in(0,1),\sigma\in(2,4)$	$c\in(-1,1),s\in(0,1)$

The SVM method is based on the LIBSVM software, and its penalty parameter is set to 2.7, and the parameter for the Gaussian kernel function is selected as 0.01. Table 4 provides further information on the values of E_{RMSE}, E_{MAE}, and R^2 for each prediction model when predicting the photovoltaic power in March.

Table 4 Power forecasting results for each model

Model	Algorithm	E_{RMSE}	E_{MAE}	R^2
SVM		0.190 46	0.148 69	0.685 89
TSK FLS (A1-C0)	BP[25]	0.245 48	0.164 26	0.623 42
	GA[26]	0.252 79	0.154 26	0.612 75
	DE[25]	0.195 97	0.115 43	0.723 91
	PSO[27]	0.202 21	0.114 05	0.765 58
	BBO	0.164 88	0.089 48	0.839 02
	GWO[18]	0.145 22	0.089 12	0.875 29
	IGWO	**0.139 76**	**0.089 10**	**0.892 64**
IT2 TSK FLS (A1-C1)	BP[14]	0.186 48	0.142 11	0.702 04
	GA[15]	0.194 72	0.131 22	0.732 63
	DE	0.208 88	0.137 08	0.736 13
	PSO	0.170 62	0.104 59	0.789 31
	BBO	0.125 07	0.088 15	0.807 98
	GWO	0.122 77	0.080 23	0.877 45
	IGWO	**0.080 13**	**0.052 08**	**0.896 33**
IT2 TSK FLS (A2-C0)	BP[14]	0.155 72	0.166 21	0.680 15
	GA[15]	0.165 55	0.160 12	0.686 22
	DE	0.187 42	0.118 90	0.712 73
	PSO	0.171 29	0.111 46	0.709 23
	BBO	0.170 29	0.106 15	0.785 74
	GWO	0.208 14	0.115 81	0.766 37
	IGWO	**0.060 93**	**0.043 39**	**0.961 76**
IT2 TSK FLS (A2-C1)	BP	0.151 93	0.111 12	0.763 42
	GA[15]	0.098 03	0.068 76	0.880 00
	DE	0.125 74	0.092 99	0.846 17
	PSO[28]	0.093 18	0.059 97	0.876 76
	BBO	0.080 32	0.052 28	0.915 85
	GWO[29]	0.097 90	0.066 01	0.930 92
	IGWO	**0.060 47**	**0.043 04**	**0.961 76**

For short-term photovoltaic power prediction one hour in advance, the proposed TSK FLS model based on IGWO optimization significantly outperforms the

other comparative models, with the A2-C1 type TSK FLS model having the lowest E_{RMSE} and E_{MAE}.

For better visualization, the boxplots of the prediction errors on the March test set for different methods are shown in Fig. 6. The visual results of the IGWO optimized A1-C0, A1-C1, A2-C0, and A2-C1 models are shown in Fig. 7. It can be seen that the TSK FLS method based on IGWO algorithm optimization has better prediction performance compared to other methods, with the A2-C1 type TSK FLS method having the best prediction performance.

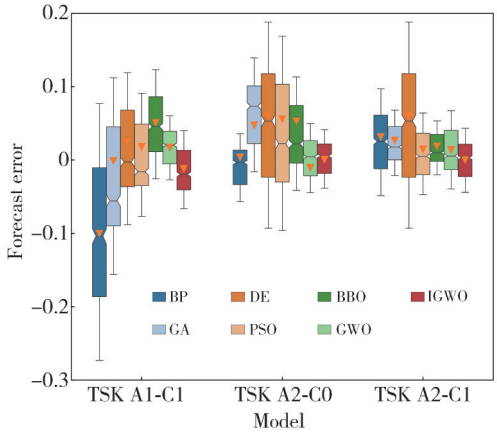


Fig. 6 Box diagram of forecasting error

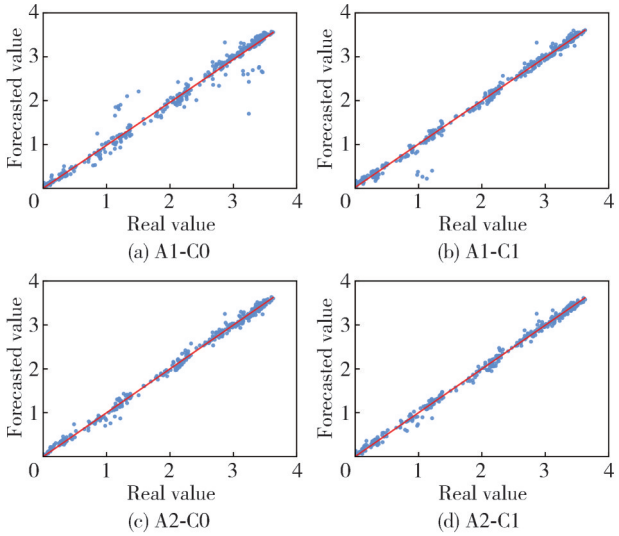


Fig. 7 Forecasting results of IT2 TSK FLS method based on IGWO optimization

To measure the optimization performance of the IGWO algorithm, taking the first type sub-data of March as an example, Fig.8 shows the convergence curve of the training error with the evolution generation when different optimization algorithms optimize the A2-C1 type TSK FLS model during the training process. Obviously, the model optimized based on the IGWO algorithm tends to be stable after about 25 iterations, and its error convergence speed is faster.

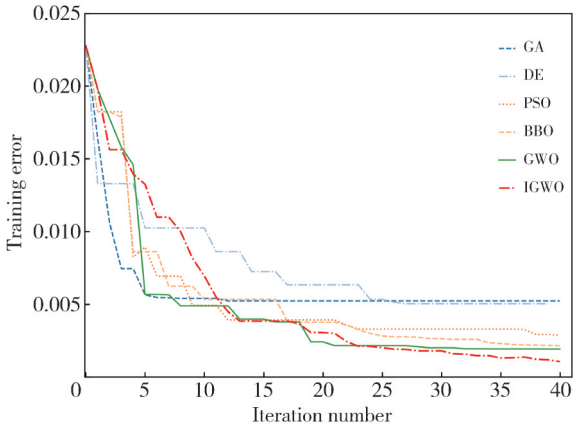


Fig. 8 Convergence of A2-C1 type TSK FLS method optimized by different algorithms

In order to observe the parameter optimization process of the fuzzy front-end membership functions (MF) by the IGWO algorithm more intuitively based on the first type sub-data of March, Fig.9 (a) — (g) respectively show the changes of the MFs for each input variable of the first rule of the A2-C1 model before and after optimization. Correspondingly, the optimization results of the posterior parameters of the five rules of the model are shown in Table 5.

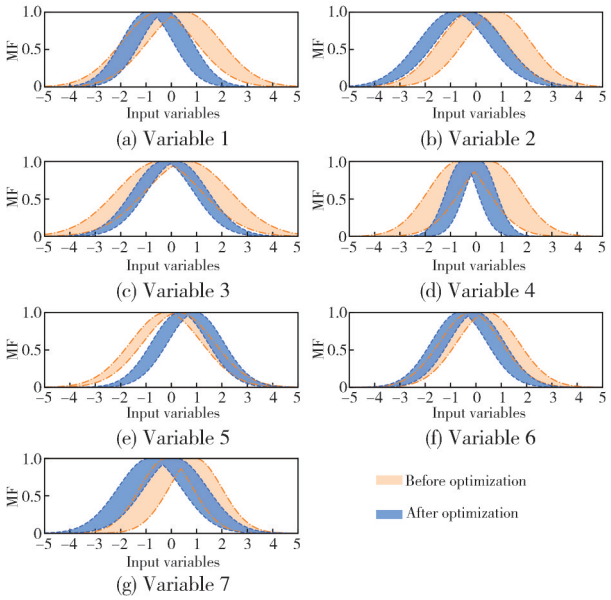


Fig. 9 Footprint of uncertainly (FOU) of each MF before and after optimization for the first fuzzy rule

To further validate the performance of the proposed model and consider the influence of seasonal factors, the IGWO optimized TSK FLS model was used for short-term PV power prediction in different seasons, and the results were compared. Data from April in autumn 2017, July in winter 2017, October in spring 2017, and January in summer 2018 were selected as the training dataset, and the corresponding monthly data from 2018 to 2019 were used for prediction analysis.

Table 6 shows the values of E_{RMSE}, E_{MAE}, and R^2 for each prediction model when predicting PV power in different seasons. It can be seen that for short-term PV power prediction one hour in advance, the IGWO optimized A2-C1 TSK FLS model has the lowest E_{RMSE}, E_{MAE}, and its prediction is significantly better than that of other models.

Table 5 Results of consequent parameter design for A2-C1 type TSK FLS method based on IGWO algorithm in PV power forecasting

Number	$y=C_0+C_1x_1+C_2x_2+C_3x_3+C_4x_4+C_5x_5+C_6x_6+C_7x_7, C=[c-s, c+s]$
Rules 1	$y_1=[-1.832\,8, 2.126\,4]+[-1.301\,2, -0.499\,7]x_1+[-0.292\,6, 0.594\,2]x_2+[-0.814\,6, 1.023\,5]x_3+[0.366\,4, 0.212\,7]x_4+[0.028\,4, 0.829\,1]x_5+[0.403\,1, 0.325\,5]x_6+[-0.070\,4, 0.073\,7]x_7$
Rules 2	$y_2=[0.044\,1, 0.705\,6]+[-0.499\,5, -0.116\,8]x_1+[1.214\,9, 0.584\,7]x_2+[-0.240\,6, 1.082\,3]x_3+[-0.363\,5, -0.026\,9]x_4+[0.716\,0, 2.577\,7]x_5+[-0.130\,4, -0.082\,2]x_6+[-0.579\,2, 0.394\,7]x_7$
Rules 3	$y_3=[0.045\,6, -0.060\,2]+[0.510\,2, 1.574\,1]x_1+[-0.334\,9, 0.836\,1]x_2+[0.285\,6, -0.549\,8]x_3+[-0.593\,1, 0.768\,0]x_4+[-0.414\,1, 0.372\,3]x_5+[-0.056\,4, 1.579\,7]x_6+[-0.303\,1, -0.468\,4]x_7$
Rules 4	$y_4=[-0.433\,1, -0.489\,1]+[0.080\,8, 0.494\,3]x_1+[-0.831\,3, -0.466\,4]x_2+[-1.383\,2, 1.573\,7]x_3+[0.282\,6, 2.324\,5]x_4+[-0.414\,4, -0.085\,4]x_5+[1.710\,1, 1.691\,9]x_6+[-0.954\,9, 1.391\,9]x_7$
Rules 5	$y_5=[-0.866\,8, 1.096\,4]+[-0.393\,7, 0.125\,2]x_1+[-0.526\,8, 0.196\,3]x_2+[-2.306\,9, 1.431\,8]x_3+[-1.068\,1, 2.093\,9]x_4+[-0.700\,4, -0.776\,7]x_5+[0.024\,3, 0.0900]x_6+[-0.105\,1, 0.758\,1]x_7$

Table 6 Power forecasting results for each model in different seasons

Season	Model	E_{RMSE}	E_{MAE}	R^2
Spring	SVM	0.130 93	0.105 92	0.763 92
	A1-C0	0.083 43	0.064 46	0.938 81
	A1-C1	0.084 89	0.066 07	0.935 46
	A2-C0	0.079 78	0.064 03	0.943 40
	A2-C1	**0.077 01**	**0.058 83**	**0.946 46**
Summer	SVM	0.109 21	0.093 77	0.814 03
	A1-C0	0.107 74	0.086 16	0.892 14
	A1-C1	0.105 06	0.083 05	0.893 95
	A2-C0	0.112 72	0.094 13	0.882 64
	A2-C1	**0.054 48**	**0.041 08**	**0.964 60**
Autumn	SVM	0.082 04	0.067 38	0.879 89
	A1-C0	0.078 19	0.060 48	0.936 31
	A1-C1	0.062 04	0.046 41	0.950 34
	A2-C0	0.060 76	0.048 13	0.961 50
	A2-C1	**0.057 55**	**0.044 74**	**0.963 57**
Winter	SVM	0.117 69	0.099 05	0.752 68
	A1-C0	0.114 89	0.098 57	0.796 49
	A1-C1	0.107 61	0.092 84	0.827 38
	A2-C0	0.106 74	0.093 41	0.827 08
	A2-C1	**0.086 03**	**0.067 69**	**0.892 45**

To better visualize the prediction performance in different seasons, Fig.10 further presents the predicted curves of each model on the test dataset. It can be seen that the predicted output curve of the IGWO optimized IT2 TSK FLS method almost completely overlaps with the curve of the system's true output, and the difference between them can only be distinguished at a very small scale, with the A2-C1 method having the best performance.

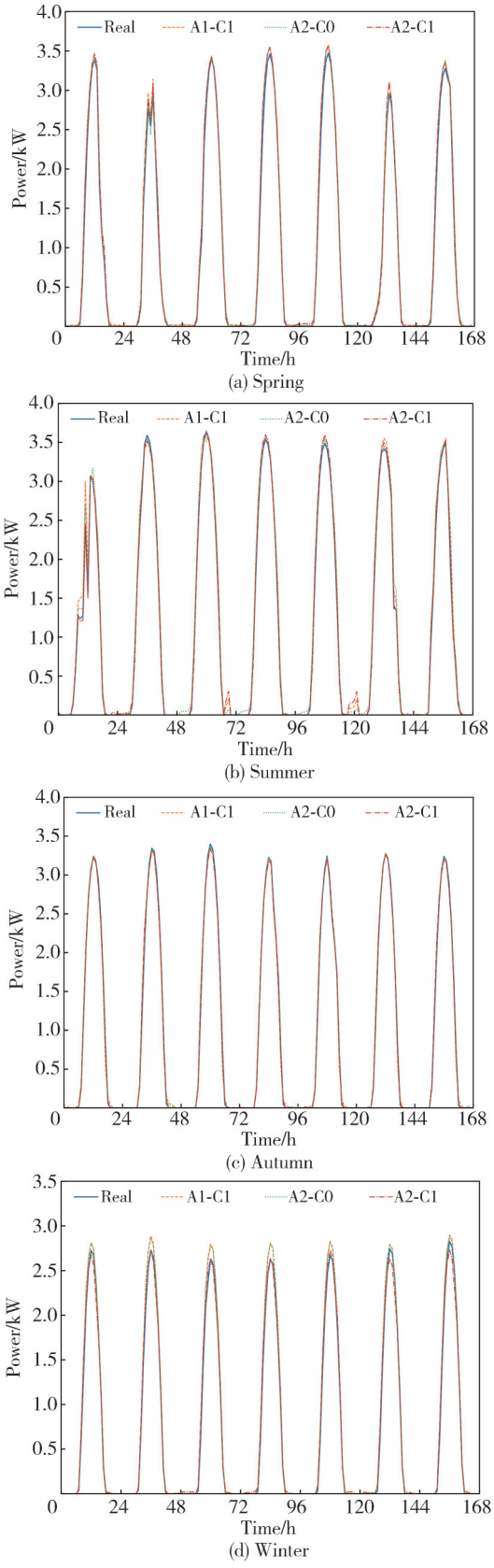


Fig. 10 Photovoltaic power forecasting curves for each model in different seasons

5 Conclusions

A method for short-term prediction of photovoltaic power was proposed by optimizing the parameters of the TSK FLS using the IGWO algorithm, which was applied to a prediction example in a certain region.

The IGWO algorithm retained the advantages of the GWO algorithm, while overcoming the its drawbacks by introducing the early maturity convergence judgment mechanism, Levy flight strategy, and nonlinear adjustment strategy of convergence factor. Meanwhile, the IT2 TSK FLS method could handle uncertainty well in modeling. Therefore, the IGWO-IT2 TSK FLS method could effectively improve the modeling accuracy, and the experimental results demonstrated the effectiveness of this method. Compared with the existing type-I TSK FLS method, the prediction accuracy of IT2 TSK FLS method has been significantly improved.

The IT2 TSK FLS method was suitable for modeling uncertain data with significant fluctuations and randomness in photovoltaic power data and had effective application potential.

Acknowledgement

This work was supported by National Natural Science Foundation of China (No. 12172157), Key Project of Natural Science Foundation of Gansu Province (No.25JRRA150), and Key Research and Development Planning Project of Gansu Province (No.23YFWA0007).

Declaration of conflicting interests

The authors have no conflict of interests related to this publication.

References

[1] WANG G, ZHANG Z, LIN J. Multi-energy complementary power systems based on solar energy: a review. Renewable and Sustainable Energy Reviews, 2024, 199: 114464.

[2] DAS U K, TEY K S, SEYEDMAHMOUDIAN M, et al. Forecasting of photovoltaic power generation and model optimization: a review. Renewable and Sustainable Energy Reviews, 2018, 81: 912-928.

[3] ÇEVIK BEKTAŞ S, ALTAŞ I H. DWT-BILSTM-based models for day-ahead hourly global horizontal solar irradiance forecasting. Neural Computing and Applications, 2024, 36(21): 13243-13253.

[4] AL-DAHIDI S, MADHIARASAN M, AL-GHUSSAIN L, et al. Forecasting solar photovoltaic power production: a comprehensive review and innovative data-driven modeling framework. Energies, 2024, 17(16): 4145.

[5] TAHIR M F, YOUSAF M Z, TZES A, et al. Enhanced solar photovoltaic power prediction using diverse machine learning algorithms with hyperparameter optimization. Renewable and Sustainable Energy Reviews, 2024, 200: 114581.

[6] KUMAR P M, SARAVANAKUMAR R, KARTHICK A, et al. Artificial neural network-based output power prediction of grid-connected semitransparent photovoltaic system. Environmental Science and Pollution Research, 2022, 29(7): 10173-10182.

[7] BEHERA M K, NAYAK N. A comparative study on short-term PV power forecasting using decomposition based optimized extreme learning machine algorithm. Engineering Science and Technology, 2020, 23(1): 156-167.

[8] LI L L, WEN S Y, TSENG M L, et al. Renewable energy prediction: a novel short-term prediction model of photovoltaic output power. Journal of Cleaner Production, 2019, 228: 359-375.

[9] SFETSOS A. A comparison of various forecasting techniques applied to mean hourly wind speed time series. Renewable Energy, 2000, 21(1): 23-35.

[10] ANSARI M, OTHMAN F, EL-SHAFIE A. Optimized fuzzy inference system to enhance prediction accuracy for influent characteristics of a sewage treatment plant. Science of the Total Environment, 2020, 722: 137878.

[11] LIANG Q L, MENDEL J M. An introduction to type-2 TSK fuzzy logic systems//FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems, August 22-25, 1999, Seoul, South Korea. New York: IEEE, 1999: 1534-1539.

[12] WU D R, ZENG Z G, MO H, et al. Interval type-2 fuzzy sets and systems: overview and outlook. Acta Automatica Sinica, 2020, 46(8): 1539-1556.

[13] MITTAL K, JAIN A, VAISLA K S, et al. A comprehensive review on type 2 fuzzy logic applications: Past, present and future. Engineering Applications of Artificial Intelligence, 2020, 95: 103916.

[14] JAFARZADEH S, FADALI M S, EVRENOSOGLU C Y. Solar power prediction using interval type-2 TSK modeling. IEEE Transactions on Sustainable Energy, 2013, 4(2): 333-339.

[15] KHOSRAVI A, NAHAVANDI S, CREIGHTON D, et al. Interval type-2 fuzzy logic systems for load forecasting: a comparative study. IEEE Transactions on Power Systems, 2012, 27(3): 1274-1282.

[16] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer. Advances in Engineering Software, 2014, 69: 46-61.

[17] MIRJALILI S. How effective is the Grey Wolf optimizer in training multi-layer perceptrons. Applied Intelligence, 2015, 43(1): 150-161.

[18] TRIPATHI S, SHRIVASTAVA A, JANA K C. Self-Tuning fuzzy controller for Sun-tracker system using Gray Wolf Optimization (GWO) technique. ISA Transactions, 2020, 101: 50-59.

[19] CHEN M, CHEN Y, XU M Z, et al. Scheduling optimization of stereo garage based on improved grey wolf optimization algorithm. Journal of North University of China (Natural Science Edition), 2023, 44(1): 58-64.
[20] HE S M, YUAN Z Y, LEI J Y, et al. Optimal setting method of inverse time over-current protection for a distribution network based on the improved grey wolf optimization. Power System Protection and Control, 2021, 49(18): 173-181.
[21] LIANG Q L, MENDEL J M. MPEG VBR video traffic modeling and classification using fuzzy technique. IEEE Transactions on Fuzzy Systems, 2001, 9(1): 183-193.
[22] DONG F F, LIU D C, WU J, et al. A method of constructing core backbone grid based on improved BBO optimization algorithm and survivability of power grid. Proceedings of the CSEE, 2014, 34(16): 2659-2667.
[23] AMIRSADRI S, MOUSAVIRAD S J, EBRAHIMPOUR-KOMLEH H. A Levy flight-based grey wolf optimizer combined with back-propagation algorithm for neural network training. Neural Computing and Applications, 2018, 30(12): 3707-3720.
[24] Desert Knowledge Australia Centre. Download Data: eco-Kinetics, 26.5 kW, mono-Si, Dual, 2010, Alice Springs. http://dkasolarcentre.com.au/historical-data/download, date accessed: 20/08/2022.
[25] KARKEVANDI-TALKHOONCHEH A, HAJIREZAIE S, HEMMATI-SARAPARDEH A, et al. Application of adaptive neuro fuzzy interface system optimized with evolutionary algorithms for modeling CO_2-crude oil minimum miscibility pressure. Fuel, 2017, 205: 34-45.
[26] YADAV H K, PAL Y, TRIPATHI M M. A novel GA-ANFIS hybrid model for short-term solar PV power forecasting in Indian electricity market. Journal of Information and Optimization Sciences, 2019, 40(2): 377-395.
[27] KALOOP M R, BARDHAN A, KARDANI N, et al. Novel application of adaptive swarm intelligence techniques coupled with adaptive network-based fuzzy inference system in predicting photovoltaic power. Renewable and Sustainable Energy Reviews, 2021, 148: 111315.
[28] YEH C Y, JENG W R, LEE S J. Data-based system modeling using a type-2 fuzzy neural network with a hybrid learning algorithm. IEEE Transactions on Neural Networks, 2011, 22(12): 2296-2309.
[29] MAO W L, HUNG C W. Type-2 fuzzy neural network using grey wolf optimizer learning algorithm for nonlinear system identification. Microsystem Technologies, 2018, 24(10): 4075-4088.

基于IGWO算法的区间二型TSK FLS方法在短期光伏功率预测中的应用

李　军*, 曾钰翔

兰州交通大学 自动化与电气工程学院, 甘肃 兰州 730070

摘　要： 针对短期光伏功率预测，在区间二型TSK模糊系统(Interval type-2 Takagi-Sugeno-Kang fuzzy logic systems，IT2 TSK FLS)的基础上，结合改进灰狼优化(Improved grey wolf optimizer，IGWO)算法策略，提出一种IGWO-IT2 TSK FLS方法。与一型TSK模糊逻辑系统方法相比，IT2 TSK FLS方法可以同时建模个体内不确定性和个体间的不确定性，在现有误差反向传播(Back propagation，BP)算法训练的基础上，进一步将IGWO算法用于模型前件参数和后件参数的设计，以进一步提高模型的预测性能。通过对灰狼优化算法进行改进，引入早熟收敛判断机制、非线性余弦调整策略、Levy飞行策略，以提高算法的收敛速度，避免局部最优。将基于IGWO优化算法的IT2 TSK FLS方法应用于某地区的光伏功率预测实例中，在同等条件下，还与一型TSK FLS方法、基于BP算法、遗传算法、差分进化、粒子群优化、生物地理学优化、灰狼优化算法等不同IT2 TSK FLS方法进行比较。实验结果表明，所提出的IGWO-IT2 TSK FLS方法在性能上优于其他方法，显示出其有效性和应用潜力。

关键词： 光伏功率；区间二型模糊逻辑系统；灰狼优化算法；模型预测性能

引用格式： LI Jun, ZENG Yuxiang. Application of interval type-2 TSK FLS method based on IGWO algorithm in short-term photovoltaic power forecasting. Journal of Measurement Science and Instrumentation, 2025, 16(2): 258-271. DOI: 10.62756/jmsi.1674-8042.2025025