



Short-term natural gas consumption prediction based on Volterra adaptive filter and improved whale optimization algorithm[☆]



Weibiao Qiao ^{a,b,*}, Zhe Yang ^c, Zhangyang Kang ^a, Zhen Pan ^d

^a School of Environmental and Municipal Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450046, China

^b Sinopec Petroleum Engineering Zhongyuan Co., Ltd, Zhengzhou 450046, China

^c School of Computer Science, The University of Manchester, Manchester, M13 9PL, UK

^d College of Petroleum and Natural Gas Engineering, Liaoning Shihua University, Fushun 113001, China

ARTICLE INFO

Keywords:

Short-term natural gas consumption
Chaotic character recognition
Phase space reconstruction
Volterra adaptive filter
Improved whale optimization algorithm
Forecast

ABSTRACT

Short-term natural gas consumption prediction is an important indicator of natural gas pipeline network planning and design, which is of great significance. The purpose of this study is to propose a novel hybrid forecast model in view of the Volterra adaptive filter and an improved whale optimization algorithm to predict the short-term natural gas consumption. Firstly, Gauss smoothing and C-C method is adopted to pretreat and reconstruct short-term natural gas consumption time series; secondly, to improve the performance of whale optimization algorithm, adaptive search-surround mechanism and spiral position and jumping behavior are introduced into it; Thirdly, Volterra adaptive filter is used to predict the short-term natural gas consumption, and the important parameters (e.g. embedding dimension) is optimized by improved whale optimization algorithm. Finally, an actual example is given to test the performance of the developed prediction model. The results indicate that (1) short-term natural gas consumption time series has chaotic characteristics; (2) performance of the improved whale optimization algorithm is better than some comparative algorithms (i.e. cuckoo optimization algorithm, etc.) based on the different evaluation indicators; (3) exploration factor is the main operational factor; (4) the performance of the proposed prediction model is better than some advanced prediction models (e.g. back propagation neural network). It can be concluded that such an innovative hybrid prediction model may provide a reference for natural gas companies to achieve intelligent scheduling.

1. Introduction

Robust prediction is the fundamental process of the energy system planning and design, as well as the foundation for natural gas companies to realize intelligent dispatching for natural gas companies (Du et al., 2017; Wang et al., 2018; Wei et al., 2019a,b,c,d; Du et al., 2019; Qiao et al., 2020, 2019a).

The short-term natural gas consumption forecasting can provide a reference for the “gas supply and sales contract” (Khotanzad and Elragal, 1999). It is the basis for pipeline network design optimization, gas storage facility design, engineering technology analysis and economic dispatching (Khotanzad et al., 2000; Hippert et al., 2001). At the same time, it is also an important indicator for realizing the modernization of natural gas pipeline network management (Viet and Mandziuk, 2003). So, the short-term natural gas consumption prediction is an important issue with obvious social and economic benefits.

The literature on the prediction of short-term natural gas consumption is more limited (Suganthi and Samuel, 2012). The utilization

of natural gas in power generation is rather limited during the last decades.

Besides, in the past few decades, oil and coal have been the main sources of energy systems. In recent years, natural gas has started to develop, which is based on its and large reserves. Besides, compared with the cost of investing in oil and coal industry, the cost of investing in natural gas industry is lower, which provides opportunities for some enterprises and leads to the rapid development of natural gas industry (Soldo, 2012). Furthermore, compared with coal and oil, natural gas burns fully and produces a small number of harmful substances. With the increasing demands on the environment, clean energy natural gas has attracted more and more attention (Deetman et al., 2013). The utilization of natural gas mainly includes industry, commerce, and civil use. Some large industries are gradually changing from coal and liquefied petroleum gas to natural gas. Commercial and civil users are also gradually changing from liquefied petroleum gas and coal to

[☆] No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to <https://doi.org/10.1016/j.engappai.2019.103323>.

* Correspondence to: School of Environmental and Municipal Engineering, North China University of Water Resources and Electric Power, No. 136 Jinshui East Road, Zhengzhou, Henan, China.

E-mail address: kaishirenheng@sina.com (W. Qiao).

natural gas, such as the rural coal to gas proposed by the Chinese government. (Ivezic, 2006). With the rapid development and full utilization of natural gas, some large enterprises have invested in the natural gas market, forming national, provincial, municipal, county and town natural gas pipelines. The development of these pipelines cannot be separated from the natural gas consumption prediction.

In recent years, different natural gas consumption forecasting models have been put forward continuously. These prediction models can be divided into two categories: the first is traditional model, and the second is intelligent model. For the traditional models, they explain the relationship between exogenous variables and short-term natural gas consumption. And the relationship is usually linear. Many traditional models are adopted to predict natural gas consumption, including ARX model (Soldo et al., 2014), SARIMAX model (Taspinar et al., 2013), ARIMAX model (Brabec et al., 2008), the hybrid forecast model (Ervural et al., 2016) and Bayesian estimation model (Sabo et al., 2011). Although these traditional models are easy to apply, there is still the shortcomings that they are in view of the linearly. They are not only difficult to deal with gas consumption with complex nonlinear characteristics but also difficult to achieve satisfactory prediction accuracy.

Based on superior nonlinear processing ability, the intelligent models have been widely applied to get higher and more satisfying prediction accuracy, such as fuzzy theory model (Azadeh et al., 2010), support vector machine model (Zhang et al., 2010; Potočnik et al., 2014; Zhu et al., 2015; Bai and Li, 2016), and neural network model (Kizilaslan and Karlik, 2008; Zhou et al., 2011; Demirel et al., 2012; Azari et al., 2012; Shi, 2013).

A short-term natural gas consumption prediction model which is the combination of multiple neural networks and multiwavelet transform is proposed (Liu et al., 2013). Compared to the combination of three neural network prediction models without multiwavelet transform and any one single neural network prediction model, the proposed prediction model is a higher inaccuracy. A coupled short-term natural gas consumption forecasting model that optimizes artificial neural networks based on genetic algorithm is developed (Karimi and Dastranj, 2014). The input variables are relative humidity, rainfall, degree-day, etc. The artificial neural networks and genetic algorithm prediction model have high prediction accuracy based on correlation coefficient etc. Authors (Yu and Xu, 2014) propose an appropriate combinational method with improved back propagation neural network and real-coded genetic algorithm that is predicted short-term natural gas consumption for Shanghai. The prediction results are based on that combinational method which outperforms several different combinational prediction models. Authors (Yu et al., 2014) propose a combinational model which is the combination of wavelet back propagation neural network and genetic algorithm to overcome the problems of the traditional back propagation neural network algorithm. Take short-term natural gas consumption from Shanghai for example, that develop the combinational model has good prediction precision. Authors (Szoplik, 2015) present the prediction results by using a multilayer perceptron model among artificial neural networks. In the process of prediction, weather, etc are as input. The prediction results indicate that for predicting short-term natural gas consumption on any hour of the day and day of the year, the multilayer perceptron model is more accurate.

These research results promote the development of natural gas consumption prediction. However, there are some problems that need to be solved in these single or coupled prediction models, such as difficult to determine the structure of neural network (Akpinar et al., 2016), low convergence speed, easy to fall into local optimal solution (Panapakidis and Dagoumas, 2017; Merkel et al., 2018) and so on. In recent years, with the development of particle swarm optimization, it has become a trend to use metaheuristic algorithms to optimize intelligent models for prediction.

In 2016, a new optimization algorithm is of easier conduction and less parameters adjusting, namely whale optimization algorithm

(WOA), proposed by Mirjalili, compared with other optimization algorithms (Mirjalili and Lewis, 2016). It only includes two main adjusted parameters (A and C). Because of its set of A , this algorithm can balance the ability of exploration and exploitation. So, it increases the possibility of local optimum. However, the search of the WOA totally depends on randomness, which leads to the algorithm's low convergence accuracy and slow convergence speed (Aljarah et al., 2018; Mafarja and Mirjalili, 2018; Mehne and Mirjalili, 2018).

In this study, a novel hybrid forecast model involving chaotic character recognition and phase space reconstruction, the widely used Volterra adaptive filter, and improved whale optimization algorithm (IWOA) proposed are successfully presented. The main contributions are summarized as follows.

(a) Development of a new metaheuristic algorithm. The IWOA is successfully developed and not only outperforms WOA and other five benchmark metaheuristic algorithms in terms of the different evaluating indicators.

(b) Determination of model input. Chaotic characteristics recognition and phase space reconstruction have been developed to dynamically determine the model input. Furthermore, IWOA is applied to optimize the three key parameters e.g. the order of the Volterra adaptive filter.

(c) Scientific and reasonable evaluation system. Model checking and four evaluation indices are introduced.

The rest of this paper is arranged as follows. Section 2 describe in detail the preprocessing methods. Section 3 give the structure of the proposed prediction model. Two experiments are showed in Section 4. The discussions are introduced in Section 5. Finally, Section 6 gives the conclusions.

2. Methods

In this section, the methods used in this study i.e. chaotic character recognition, phase space reconstruction, Volterra adaptive filter, WOA, and IWOA are described in detail as follows.

2.1. Pretreatment method

2.1.1. Gauss smoothing

The short-term natural gas consumption time series collected can be regarded as a signal, and the signal is smoothed to filter out the high-frequency part, to reduce the impact of noise. One-dimensional Gaussian smoothing (Tronarp and Särkkä, 2019) is computed by Eq. (1):

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, x \in [1, 2, \dots, n] \quad (1)$$

where x whose value is from 1 to n is the component, and σ whose value is 3 in this study is the order of one-dimensional Gaussian smoothing.

2.1.2. Chaotic character recognition

The chaotic characteristics of the dynamical system can be judged by analyzing whether the system is sensitive to the initial value. In addition, the maximum Lyapunov exponent, which is one of the important bases for recognizing the chaotic character of the system, is the characteristic quantity to characterize this characteristic. In the early years, the maximum Lyapunov exponent is mainly computed by the Wolf method (Wolf et al., 1985). After that, one new calculation method which is called small data quantity method is proposed by Rosenstein (Rosenstein et al., 1993). And its calculation process is as follows:

The key of chaotic character recognition is looking for the nearest neighbor point $\mathbf{X}(j)$ of each state point $\mathbf{X}(j)$ on an orbit, so the relationship between $\mathbf{X}(j)$ and $\mathbf{X}(j)$ is expressed by Eq. (2):

$$d_j(0) = \min_{\mathbf{X}(j)} \|\mathbf{X}(j) - \mathbf{X}(\hat{j})\|, |j - \hat{j}| \geq p \quad (2)$$

where the range of the parameter j is within range $N_0, N_0 + 1, \dots, N, N_0$ is $(m-1) + 1$; p represents the average period of the time series which is calculated by fast Fourier transform; $d_j(0)$ represents the distance between a pair of nearest neighbors point at the initial moment, and $X(j)$ is the state point in the reconstructed phase space. Sato et al. estimate that the maximum Lyapunov exponent is expressed by Eq. (3):

$$\lambda_1(i) = \frac{1}{i \cdot \Delta t} \cdot \frac{1}{N-i} \sum_{j=1}^{N-i} \ln \frac{d_j(i)}{d_j(0)} \quad (3)$$

where the range of the parameter j which indicates the sample data cycle is $N_0, N_0 + 1, \dots, N$, $d_j(i)$ is the distance. The parameter C_j is expressed as follows:

$$C_j = d_j(0) \lambda_1(i, k) = \frac{1}{k \cdot \Delta t} \cdot \frac{1}{N-k} \sum_{j=1}^{N-k} \ln \frac{d_j(i+k)}{d_j(i)} \quad (4)$$

where k is a constant. In combination with Sato's estimation, the parameter $d_j(i)$ is calculated by Eqs. (5) and (6):

$$d_j(i) \approx C_j e^{\lambda_1(i \cdot \Delta t)} \quad (5)$$

$$C_j = d_j(0) \quad (6)$$

Eqs. (5) and (6) are taken reciprocal of both sides, which are expressed by Eq. (7):

$$\ln d_j(i) \approx \ln C_j + \lambda_1(i \cdot \Delta t) \quad (7)$$

The maximum Lyapunov exponent can be approximated as a set of linear slopes represented by the Eq. (7) and can be obtained through using the least squares method to approximate this set of lines and it is expressed by Eq. (8):

$$y(i) = \frac{1}{w \Delta t} \sum_{j=1}^w \ln [d_j(i)] \quad (8)$$

where w is the number of nonzero $d_j(i)$, $y(i)$ is the average of the distance $d_j(i)$ to w cumulative sums.

2.1.3. Phase space reconstruction

The purpose of phase space reconstruction is to reconstruct low-dimensional space into high-dimensional space, to solve the problem. The solution to the problem is to find the optimal delay time τ and embedding dimension m . In this subsection, the C-C method (Cao, 1997) is applied to determine the optimal delay time and the embedding dimension.

The chaotic time series is $\{x(i), i = 1, 2, \dots, N\}$, m is the embedding dimension, τ is the optimal delay time, $X_j(i) = \{x_j(i), x_j(i+\tau), \dots, x_j(i+(m-1)\tau)\}$ ($i = 1, 2, \dots, N$) is the point in the phase space, thus the time integral of short-term natural gas consumption is defined by Eq. (9):

$$C(m, N, r, \tau) = \frac{1}{M} \sum_{1 \leq j \leq k \leq M} \theta \left(r - \|X_j - X_k\| \right) \quad (9)$$

where N is the length, r is the neighborhood radius, $\theta(\cdot)$ is a Heaviside unit function and it is expressed by Eq. (10):

$$\theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \quad (10)$$

Associative dimension is expressed as follows:

$$D(m, \tau) = \lim_{r \rightarrow 0} \frac{\log C(m, N, r, \tau)}{\log r} \quad (11)$$

where $\log C(m, r, \tau)$ is equal to the limit of $C(m, N, r, \tau)$ and $\log C(m, r, \tau)$ represents is less than r . Definition of the statistics of $\{x(i)\}$ is expressed by Eq. (12):

$$S(m, N, r, \tau) = C(m, N, r, \tau) - C^m(m, N, r, \tau) \quad (12)$$

where $S(m, N, r, \tau)$ reflects the autocorrelation properties of the sequence, and τ is the first zero point of $S(m, N, r, \tau)$ or the time point of

the smallest difference between all the radius r . The difference between the maximum and minimum of two radius is expressed by Eq. (13):

$$\Delta S(m, \tau) = \max [S(m, N, r_j, \tau)] - \min [S(m, N, r_k, \tau)] \quad (13)$$

where the parameter j is not equal to the parameter k ; $\Delta S(m, \tau)$ measures the maximum deviation for all radius. According to the statistical principle, the r value is between 2 and 5. The equation is expressed by Eqs. (14) → (16):

$$S_{cor}(\tau) = \Delta \bar{S}(\tau) + |\bar{S}(\tau)| \quad (14)$$

$$\bar{S}(\tau) = \frac{1}{n_j n_m} \sum_{j=2}^4 \sum_{m=1}^5 S(m, N, r, \tau) \quad (15)$$

$$\Delta \bar{S}(\tau) = \frac{1}{n_m} \sum_{m=2}^5 \Delta S(m, N, r, \tau) \quad (16)$$

where the parameter n_j and n_m are 4 and 4, respectively.

2.2. Volterra adaptive filter

A preliminary framework of nonlinear adaptive chaotic prediction techniques (Kashani et al., 2016) for chaotic time series is established, which has been widely used at present. Volterra adaptive filter is applied in this study. The basic theory of the Volterra adaptive filter is described as follows.

Assuming that the input vector is $x(t)$ ($t = 1, 2, \dots, M$) and the output vector is $Y(t) = \hat{x}(t+1)$ ($t = 1, 2, \dots, M$), the Volterra series expansion of the nonlinear system is expressed by Eq. (17):

$$\begin{aligned} \hat{x}(t+1) = F[x(t)] &= h_0 + \sum_{m=0}^{+\infty} h_1(m)x(t-m) \\ &+ \sum_{m_1=0}^{+\infty} \sum_{m_2=0}^{+\infty} h_2(m_1, m_2)x(t-m_1)x(t-m_2) + \dots \\ &+ \sum_{m_1=0}^{+\infty} \sum_{m_2=0}^{+\infty} \dots \sum_{m_p=0}^{+\infty} \left[h_p(m_1, m_2, \dots, m_p)x(t-m_1) \right. \\ &\quad \left. \times x(t-m_2) \dots x(t-m_p) + \dots \right] \end{aligned} \quad (17)$$

where h_p is p order Volterra kernel, and m is the input dimension of the filter.

Volterra infinite series expansion is difficult to achieve in practice, and it must take the limited cut and limited summation form. Take the second-order truncation as an example, the form is expressed as follows:

$$\begin{aligned} \hat{x}(t+1) = F[x(t)] &= h_0 + \sum_{m=0}^{N_1} h_1(m)x(t-m) + \sum_{m_1=0}^{N_2-1} \sum_{m_2=0}^{N_2-1} h_2(m_1, m_2) \\ &\quad \times x(t-m_1)x(t-m_2) \end{aligned} \quad (18)$$

The signal of the nonlinear extension is expressed by Eq. (19):

$$U(t) = [u(t), u(t-1), \dots, u(t-L-1)]^T \quad (19)$$

where $U(t)$ is the input vector of the filter, and the parameter L is equal to or greater than the parameter N_1 and N_2 .

In application design, N_1 and N_2 are finite lengths. According to Takens's embedding theorem, we can make that the parameter N_1 is equal N_2 and m which is equal or greater than the parameter $2D_2 + 1$, D_2 represents the associated dimension.

The filter of the chaotic time series prediction is expressed by Eq. (20):

$$\hat{x}(t+1) = h_0 + \sum_{i=0}^{m-1} h_1(i)x(t-i) + \sum_{i=0}^{m-1} \sum_{j=i}^{m-1} h_2(i, j)x(t-i) \times x(t-j) \quad (20)$$

The total number of coefficients after the state expansion is $L = 1 + m + m(m+1)/2$. The input vector $U(t)$ of the linear adaptive finite

impulse filter is defined by Eq. (21):

$$\mathbf{U}(t) = [1, x(t), x(t-1), \dots, x(t-m+1), x^2(t), x(t)x(t-1), \dots, x^2(t-m+1)]^T \quad (21)$$

The coefficient vector is expressed by Eq. (22):

$$\mathbf{H}(t) = [h_0, h(0), h(1), \dots, h(m-1), h_2(0,0), h_2(0,1), \dots, h_2(0,m-1)]^T \quad (22)$$

Based on Eqs. (21) and (22), Eq. (20) can be expressed as:

$$\hat{x}(t+1) = \mathbf{H}^T(n)\mathbf{U}(n) \quad (23)$$

As for the Volterra adaptive filter of the Eq. (23), the time orthogonal adaptive algorithm can be used to solve. For the input vector $\mathbf{U}(t)$ and the coefficient vector $\mathbf{H}(t)$, and the coefficient vector is expressed by Eq. (24):

$$\begin{aligned} x(t) &= d(t), \hat{x} = \hat{d}(t), \\ \hat{d}(t) &= \mathbf{H}^T(t-1)\mathbf{U}(t-1), \\ \mathbf{H}(t) &= \mathbf{H}(t-1) + c \times \frac{e(t-1)}{\mathbf{U}^T(t)\mathbf{U}(t)} \mathbf{U}(t-1), \\ e(t) &= d(t) - \hat{d}(t) \end{aligned} \quad (24)$$

where c is the parameter that controls the convergence.

2.3. Basic and improved whale optimization algorithm

2.3.1. Basic whale optimization algorithm

In 2016, WOA is proposed by Mirjalili, which is a new metaheuristic algorithm. In WOA, every humpback whale stands for a feasible solution. Furthermore, the process of humpback whale hunting is very special, which is called spiral bubble predation. In addition, to better understand the WOA, the detailed steps of the WOA are given as follows:

Step1: Contraction enclosure mechanism

That humpback whales will surround their prey is called the contraction enclosure mechanism. To describe the contraction enclosure mechanism, the Eqs. (25) and (26) are given as follows:

$$D = |CX^*(t) - X(t)| \quad (25)$$

$$X(t+1) = X^*(t) - AD \quad (26)$$

where t represents the current number of iterations; $X^*(t)$ represents the best whale position vector by far; $X(t)$ stands for the current position vector of the whale; A and C represent the vector coefficient, which is calculated by Eqs. (27) and (28).

$$A = 2ar_1 - a \quad (27)$$

$$C = 2r_2 \quad (28)$$

where r_1 and r_2 stand for the casual numbers in $(0, 1)$; t stands for the current number of iterations; T_{\max} stands for maximum iteration times; a represents a convergent factor whose value decreases linearly from 2 to 0, which is calculated by Eq. (29).

$$a = 2 - 2T/T_{\max} \quad (29)$$

Step2: Update spiral position

According to the hunting behavior of humpback whales, they swim toward their prey in a spiral motion, so the spiral position updating is expressed by Eq. (30):

$$X(t+1) = X^*(t) + D_p e^{bl} \cos(2\pi l) \quad (30)$$

where $D_p = |X^*(t) - X(t)|$ represents the distance between the whale and its prey; $X^*(t)$ represents the best position vector by far; b stands for a constant, and l is the casual number in $(0, 1)$.

It is worth noting that while the whale swims to its prey in a spiral like manner, it also must contract to encircle it. Therefore, the

contraction encircling mechanism is chosen by the probability P_i , and the spiral model is chosen by the probability $1 - P_i$. The calculation process is showed by Eq. (31):

$$X(t+1) = \begin{cases} X^*(t) - AD & p < P_i \\ X^*(t) + D_p e^{bl} \cos(2\pi l) & p \geq P_i \end{cases} \quad (31)$$

That it attacks prey and becomes close to prey is set on the mathematical model to reduce the value of a , so that A 's range also fell with a in the process of iteration. When a 's value drops from 2 to 0, A is within $[-a, a]$ random value. What is more, when A 's value is $[-1, 1]$, the whale's next position can be its position now and location anywhere between preys. When A is less than 1, the whale attacks on its prey.

The humpback whale swims along the spiral path while surrounding its prey. To simulate this hunting behavior of whales, the probability of the contraction enclosure mechanism is set to 0.5 and the probability of the helix position update is set to 0.5.

Step3: Random searching

When the whale is a random search for its prey, it needs to change its position by random searching. The calculation process is expressed by Eqs. (32) and (33):

$$D = |CX_{rand} - X(t)| \quad (32)$$

$$X(t+1) = X_{rand} - AD \quad (33)$$

where X_{rand} stands for the whale's position vector of casual choosing. When $A \geq 1$, through random searching random search a searching agent will renew the position of other whales according to the random searching whale, which will force the whale to get away from the prey so that it can find better prey. In this way, the ability of algorithm exploration can be strengthened to make WOA available to search from all aspects.

2.3.2. Improved whale optimization algorithm

Compared with other metaheuristic algorithms, WOA is of easier conduction and less parameters adjusting. It only includes two main adjusted parameters (A and C). Because of its set of A , this algorithm has the ability to balance exploration and exploitation. So, it increases the possibility of local optimum. However, in WOA, the A and C are set to 0.5, respectively, which is obviously unreasonable (Mafarja and Mirjalili, 2017; Abdel-Basset et al., 2018). Meanwhile, WOA's ability to search for optimization of all aspects needs to be further improved.

To solve the shortcomings of WOA, the improved details are as follows:

Step1: Adaptive search enveloping mechanism and spiral position

In WOA, the A and C are set to 0.5. However, in the early stage of the search, In the early days of whale hunting, in order to achieve a broader search, the best solution should be found as much as possible. In the late stage of whale hunting, more accurate searching should be realized, so as to ensure to find the optimal solution. The new definition of two kinds of mechanism is shown by Eqs. (34) → (36):

$$p_{t+1} = \begin{cases} p_0 & t = 1 \\ p_t * a + p_{\min} & t > 1 \end{cases} \quad (34)$$

$$p'_t = 1 - p_t \quad (35)$$

$$a = \exp \left[-30 \times \left(t/T_{\max} \right)^S \right] \quad (36)$$

where p_0 is the initial adaptive contraction surrounding mechanism probability; p_t and p_{t+1} are contraction surrounding mechanism probabilities of generation t th and $(t+1)$ th; p_{\min} stands for the minimal enveloping probability; p'_t is helix position update probability of generation t th; t stands for number of iterations; T_{\max} stands for maximum iterations and $S = 2$.

Step2: jumping behavior is introduces

When the jumping behavior of whale is introduced in this study, whale tries to break away from the region the current local optimum

Table 1

Statistical description based on collected samples.

Sample	SN	Mean. (m^3)	Max. (m^3)	Min. (m^3)	Std. (m^3)
All	720	19 924.33	40 780	5620	8601.408
Training	696	19 822.83	40 780	5620	8581.54
Test	24	22 867.92	38 470	9990	8841.003

value can fall into a local minimum, by changing the position of whale randomly.

The calculation process of the jumping behavior is expressed by Eq. (37).

$$X_i(t+1) = X_i(t) + c(1 - 2\text{rand})(\max(X_{\text{all}}) - \min(X_{\text{all}}))/2 \quad (37)$$

where c stands for the jumping coefficient and X_{all} stands for all whale.

The pseudo-code of IWOA is listed as follows:

3. The framework of the hybrid forecast model

The flowchart of the proposed hybrid prediction model in this study is depicted in Fig. 1, which consists of four modules: (A) Preprocessing module; (B) Improved whale optimization algorithm module; (C) The whole algorithm module; (D) Evaluation module. Furthermore, the detailed explanation is listed as follows:

(A) Preprocessing module: in order to determine if there is chaotic character of the short-term natural gas consumption, the small data quantity method is applied to calculate the maximum Lyapunov exponent. In addition, noise from short-term gas consumption is filtered using third-order Gaussian smoothing. Furthermore, the C-C method is applied to confirm the input forms of the proposed hybrid forecast model. Then, the normalization method is adopted.

(B) Improved whale optimization algorithm module: in view of the shortcomings of the whale optimization algorithm, two improvement strategies are proposed in this module.

(C) The whole algorithm module: combining phase space reconstruction and Volterra adaptive filter, a novel hybrid forecast model is proposed. The IWOA is developed in this study to search for the optimal parameters (m , τ , the order of Volterra adaptive filter), which is conducive to enhance the forecast accuracy.

(D) Evaluation module: four metrics which are RMSE (root mean square error), MAE (mean absolute error), MAPE (mean absolute percent error), and Pearson correlation coefficient (R^2), respectively. What is more, posttest indexes Q , C , p (Augsteijn et al., 2019) is used to evaluate the model performance.

4. Experiments results and analysis

In this section, two main research works were carried out: first, IWOA will be tested; second, the performance of the hybrid forecasting model will be tested. The experimental results and analysis of these two studies are given in detail below.

4.1. Experimental sample, design, parameter setting, evaluating indicator, comparison algorithms

4.1.1. Experimental sample and design

In this subsection, a set of samples from Hefei (E 117.27 and N 31.86) natural gas gate station in Anhui province, China is collected which is shown in Fig. 2. Furthermore, five statistical indicators including SN (sample number), Mean, Min (Minimum), Max (Maximum), and Std. (standard deviation) of sample collection are applied to implement the work. Specific statistical results are shown in Table 1.

To verify the effectiveness of the developed model, two experiments are executed: Experiment I: algorithm test and Experiment II: the performance comparison of different prediction models. In Experiment II, to test the proposed IWOA and other five metaheuristic algorithms,

the performance comparison of IWOA and other five metaheuristic algorithm is executed. Furthermore, the performance of two operational factors (exploration and exploration) of IWOA is tested. What is more, in Experiment II, the phase space of the samples is reconstructed. On this basis, the performance of different prediction models is compared.

In order to ensure the comparison fairness of all experiments, two experiments are operated in the MATLAB whose edition is R2018a, the operating environment is Windows 10 with a 64-bit 2.60 GHz Intel Core i7 6700 CPU and 32.00 GB of RAM.

4.1.2. Experimental comparison algorithm and parameter setting

To verify the performance of the developed model, cuckoo optimization algorithm (COA) (Zhu and Wang, 2019), fish optimization algorithm (FOA) (Zeng et al., 2019), improved genetic optimization algorithm (IGOA) (Malakoti-Moghadam et al., 2019), improved particle swarm optimization algorithm (IPSOA) (Mousavi et al., 2019), and WOA are used as comparison algorithm. In Experiment II, Gaussian smoothing-back propagation neural network (GS-BPNN) (Dai et al., 2019), Gaussian smoothing-general regression neural network (GS-GRNN) (Dai et al., 2019), Gaussian smoothing-Elman neural network (GS-ELMANN) (Lu et al., 2019), Gaussian smoothing-phase space reconstruction-Volterra adaptive filter (GS-PSR-VAF), Gaussian smoothing-least squares support vector machine (GS-LSSVM) (Guo et al., 2019) are applied as comparison algorithm. Based on the principle of experimental fairness, the comparison of the two experiments should be carried out under the same parameters. Furthermore, the parameters of Experiment I are set: maximum iteration is 50; the population size is 16; the number of experiments is 100; other detailed parameters are set in Table 2. What is more, the parameters of Experiment II are set as follows: the order of Gaussian smoothing is 3; the population size is 16; maximum iteration is 50; the length of the sliding time window is 10; other detailed parameters are listed in Table 2.

4.1.3. Experimental evaluating indicator

The minimum of the minimum according (MMA-1), the maximum of the minimum according (MMA-1), the average of the minimum according (AMA), the standard deviation of the minimum according (STMA), the average number of generation (AG), the average success rate (ASR) and the average execution time (AET) (Qiao and Yang, 2019b,c) are used as evaluating indicator in Experiment I.

The Q , C , p , RMSE, MAE, MAPE and R^2 (Bai et al., 2019; Wei et al., 2019a,b,c,d), AET is used as evaluating indicator in Experiment II. In addition, the detailed definitions of RMSE, MAE, MAPE, and R^2 are defined by Eqs. (38) → (40):

$$\text{RMSE} = \sqrt{\sum_{t=1}^N (\hat{y}(t) - y(t))^2 / N} \quad (38)$$

$$\text{MAE} = \sqrt{\sum_{t=1}^N |\hat{y}(t) - y(t)| / N} \quad (39)$$

$$\text{MAPE} = \left(\sum_{t=1}^N |(\hat{y}(t) - y(t)) / y(t)| \right) / N \quad (40)$$

$$R^2 = \frac{\sum_{t=1}^N (y(t) - \bar{y}(t)) (\hat{y}(t) - \bar{\hat{y}}(t))}{\sqrt{\sum_{t=1}^N (y(t) - \bar{y}(t))^2 \cdot \sum_{t=1}^N (\hat{y}(t) - \bar{\hat{y}}(t))^2}} \quad (41)$$

where N represents the number of the sample; $\hat{y}(t)$ and $y(t)$ is the forecast and actual sample; t represents the number of test and train sample; $\bar{y}(t)$ and $\bar{y}(t)$ stands for the average value of the prediction and actual samples, respectively.

Algorithm 1 (IWOA)

Input: $y_t^{(0)} = (y^{(0)}(1), y^{(0)}(2), \dots, y^{(0)}(q))$ —— the training samples
 $y_f^{(0)} = (y^{(0)}(q+1), y^{(0)}(q+2), \dots, y^{(0)}(q+l))$ —— the test samples

Output $\hat{y}_f^{(0)} = (\hat{y}^{(0)}(q+1), \hat{y}^{(0)}(q+2), \dots, \hat{y}^{(0)}(q+l))$ —— the forecast dataset

Parameters: p_{\min} — minimum probability of adaptive shrink wrapping mechanism; p_0 — initial probability of adaptive shrink wrapping mechanism; Bound — scope of X ; m — population; T — maximum iterations; b — whale algorithm constant; s — adaptive coefficient

```

1  /* Set parameters of IWOA. */
2  /* Initial population. */
3  X=initialPopulation (boundsstate, boundη, boundL, popsize)
4  /* Calculating the fitness. */
5  Fit=calculateFitness (X, Ftr, seed, m)
6  /* Update the optimal position of history individual. */
7  [X_p, Fit_p] = updataXp(X, Fit)
8  For each i in 1:T
9  	/* Update Adaptive Probability. */
10  a=exp(-30/(i/T)s)
11  If i==1
12  	Temp_p=p0
13  Else
14  	Temp_p=Temp_p*a+ pmin
15  END (If)
16  /* Update population individuals. */
17  For each j in 1: m
18  	If rand<Temp_p
19   	/*Contraction encirclement. */
20   >New_X(i)=shrink(X(i))
21  Else
22   	/*Spiral upward. */
23   >New_X(i)=spiralism(X(i))
24  END (if)
25  END (for)
26  /* To determine whether the algorithm has converged. */
27  If variation(Fit_p,Last_fit_P)< min_change && i>1
28  /*In the population, some X is selected randomly for jumping behavior. */
29  New_X=jump(New_X)
30  END (if)
31  /* Recalculate fitness. */
32  New_fit = calculateFitness(New_X, Ftr, seed, m)
33  Last_fit_p=fit_p;
34  /* Update the optimal position of history individual. */
35  [X_p, Fit_p] = updataXp(New_X, New_fit, X_p, Fit_p)
36  END (for)

```

4.2. Experiment I: Algorithm test**4.2.1. The performance comparison of IWOA and other five metaheuristic algorithm**

In this subsection, to determine the performance of IWOA and other five metaheuristic algorithm, four test functions including Griewank function, Schaffer function (SF), Rastrigin function (RF), and Ackley function (AF) (Qiao and Yang, 2019b,c; Qiao et al., 2020) are used to verify the superiority of IWOA, which are defined by Eqs. (42) → (45). Their two-dimensional diagrams are shown in Fig. 3.

The definition of RF is as follow:

$$f1(x_1, x_2) = 20 + x_1^2 + x_2^2 - 10(\cos 2\pi x_1 + \cos 2\pi x_2) \quad (42)$$

where the range of x_1 and x_2 is in $[-5, 5]$ and $[-5, 5]$; the optimal solution of $F1(x_1, x_2)$ is equal to 0.

SF is defined as follow:

$$f2(x_1, x_2) = 0.5 + (\sin^2(x_1^2 - x_2^2) - 0.5) / \left([1 + 0.001(x_1^2 + x_2^2)]^2 \right) \quad (43)$$

where the range of x_1 and x_2 is in $[-5, 5]$ and $[-5, 5]$; the optimal solution of $F2(x_1, x_2)$ is equal to 0.

AF is defined as follow:

$$f3(x) = -a \exp \left(-b \sqrt{\frac{1}{d} \sum_{i=1}^d x_i^2} \right) - \exp \left(\frac{1}{d} \sum_{i=1}^d \cos(cx_i) \right) + a + \exp(1) \quad (44)$$

where i is equal to 1 and 2; $d = 2$; $b = 0.2$; $a = 20$; $c = 2\pi$; x_1 and x_2 are belongs to $[-10, 10]$, respectively; the optimal solution of $F3(x_1, x_2)$ is 0.

GF is defined as follow:

$$f4(x) = \sum_{i=1}^N \frac{x_i^2}{4000} - \prod_{i=1}^N \cos \left(x_i / \sqrt{i} \right) + 1 \quad (45)$$

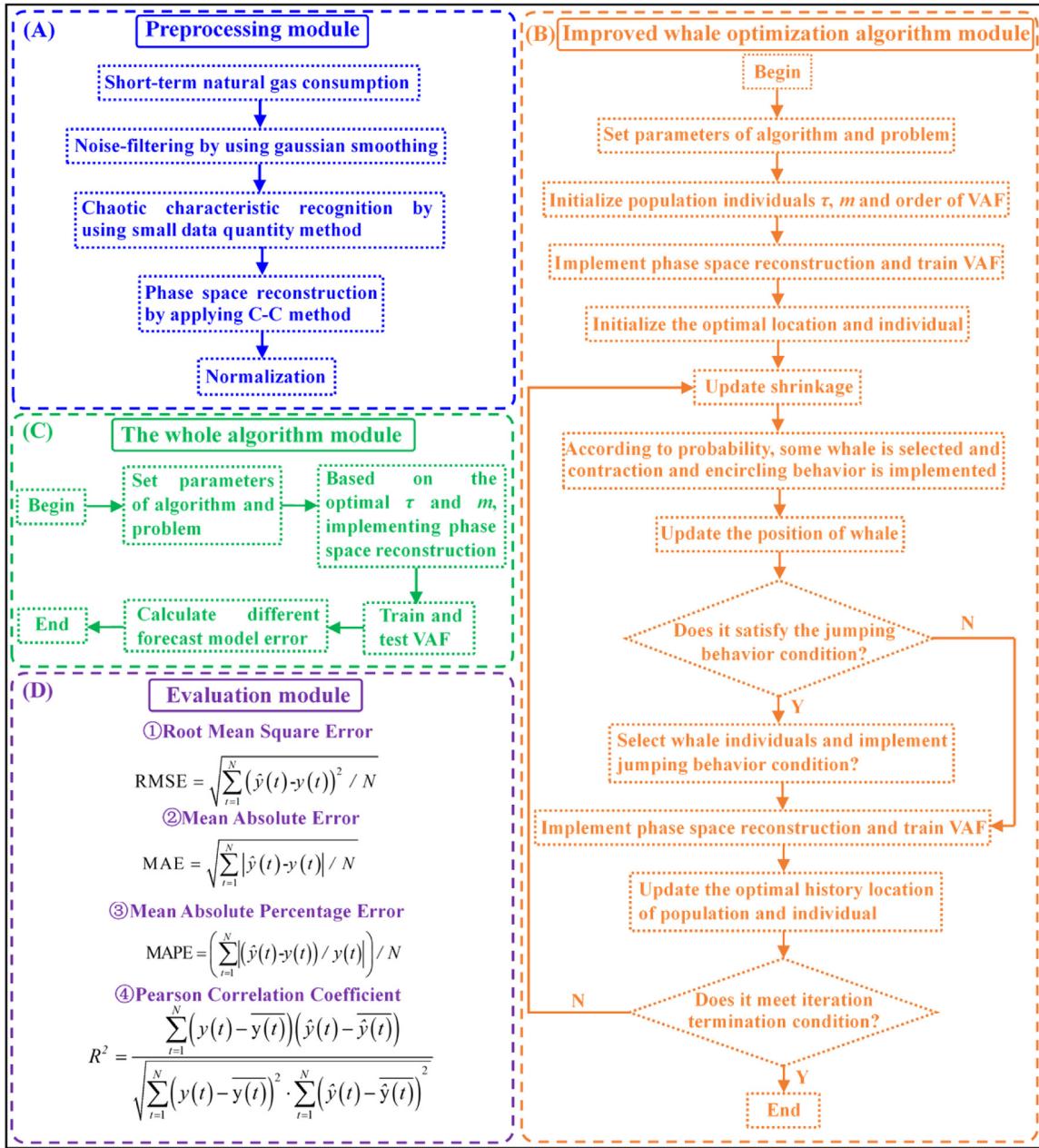


Fig. 1. The flowchart of the proposed forecast model.

where i is from 1 to 2; x_1 and x_2 are belongs to $[-10, 10]$, respectively; the optimal solution of $F4(x_1, x_2)$ is 0.

Comparison of iterative behavior by using IWOA and some metaheuristic algorithms (i.e. COA et al.) based on the four test functions are displayed in Fig. 4. Table 3 lists the performance comparison of IWOA and some metaheuristic algorithms (i.e. COA et al.) based on the four test functions. By analyzing and comparing Fig. 4 and Table 3, some key results and analysis are shown as follows:

(a) In Comparison I, for Fig. 4, the generation bight of IWOA is lower than those of the other five optimization algorithms (e.g. COA). This result indicates that IWOA is easier to converge.

(b) In Comparison II, from Table 3, for RF, IWOA is 0.61920 → 3.86956, 1.90251 → 3.85318, 0.61971 → 3.49404, and 0.28906 → 1.70503 lower than WOA, COA, FOA, IGOA, and IPSOA based on MMA-1, MMA-2, AMA, and STMA. Furthermore, IWOA is -19.81 → -0.54, and -0.50 → -0.07 higher than WOA, COA, FOA, IGOA, and IPSOA based on AG, and ASR. For SF, IWOA is 0.00074 → 0.00239, 0.00213

→ 0.01005, 0.00051 → 0.00237, and 0.00083 → 0.00265 lower than WOA, COA, FOA, IGOA, and IPSOA based on MMA-1, MMA-2, AMA, and STMA. Furthermore, IWOA is -32.89 → -3.01, and -0.77 → -0.42 higher than WOA, COA, FOA, IGOA, and IPSOA based on AG, and ASR. For AF, IWOA is 0.43343 → 3.08260, 1.19487 → 4.94189, 0.60952 → 3.35843, and 0.49577 → 1.07521 lower than WOA, COA, FOA, IGOA, and IPSOA based on MMA-1, MMA-2, AMA, and STMA. Furthermore, IWOA is -0.92 → -0.53 higher than WOA, COA, FOA, IGOA, and IPSOA based on ASR. For GF, IWOA is 0.00304 → 0.02444, 0.00035 → 0.08153, 0.00770 → 0.02813, and 0.00070 → 0.00255 lower than WOA, COA, FOA, IGOA, and IPSOA based on MMA-1, MMA-2, AMA, and STMA. Furthermore, IWOA is -29.52 → -0.12, and -0.89 → -0.22 higher than WOA, COA, FOA, IGOA, and IPSOA based on AG, and ASR.

(c) In Comparison III, for RF, the optimal solution obtained by IWOA is 0.51382, which is close to the real global optimal solution, comparing with WOA, COA, FOA, IGOA, and IPSOA. Again, for the other three test functions, the optimal solution gained by IWOA is the approach to the

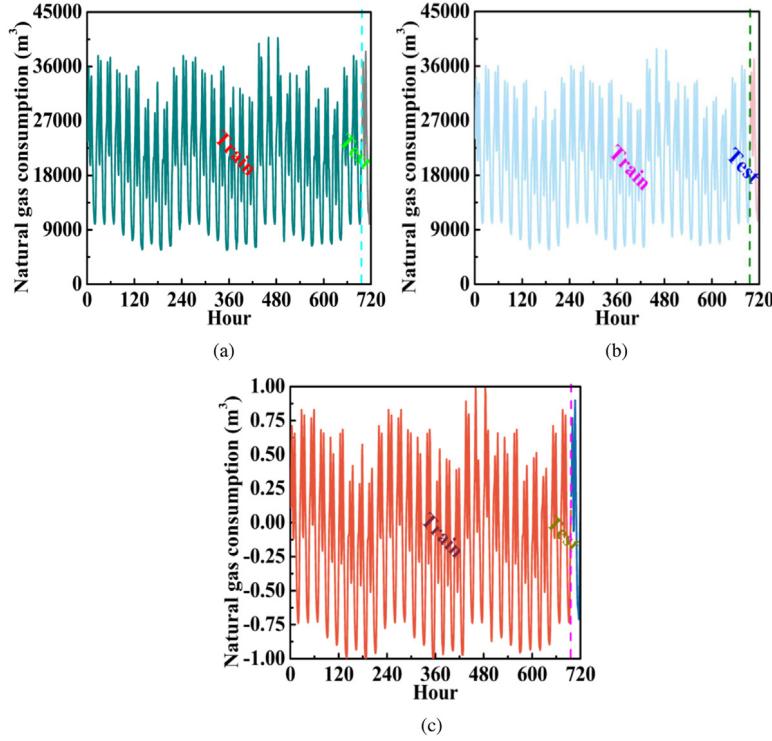


Fig. 2. Studied sample: (a) original short-term natural gas consumption; (b) natural gas consumption after Gauss smoothing; (c) normalized natural gas consumption.

Table 2
The setting of experimental parameters.

Algorithm and model	Parameter setting	Value	Parameter setting	Value
IWOA	Maximum weight	0.8	Minimum weight	0.5
	Jumping behavior probability	0.1	p	0.3
	Minimum error of jumping	0.0001		
WOA	Maximum weight	0.8	Minimum weight	0.5
	p	0.3		
COA	Step size control variable	0.2	Inertia weight	5
	Beta	2	Sigma	0.6966
FOA	View	0.2	Step	0.8
	Predation attempts	10		
IGOA	Cross probability	0.8	Mutation probability	0.1
IPSOA	Learning factor 1	1	Learning factor 2	1
	Maximum weight	0.8	Minimum weight	0.5
BPNN	Neurons structure	[5,5]	Network targets	0.001
	Initial layer delay conditions	0	Initial input delay conditions	0
GRNN	Spread of radial basis functions	1.0		
ELMANNN	Neurons structure	10	Network targets	0.001
	Initial layer delay conditions	0	Initial input delay conditions	0
LSSVM	σ	0.1	γ	1

global solution. Especially, the gap between the optimal solution gained by IWOA and the global solution is less than 0.00050.

(d) In Comparison IV, IPSOA and WOA converge when the number of iterations is less than 20, and the optimal solutions obtained by IPSOA and WOA for the four test functions are quite different from the global solutions, which can be considered as falling into the local minimum and converging prematurely. However, the optimal solution obtained by IWOA has a smaller difference from the real global optimal, which can be considered as not falling into the local minimum and jumping out successfully. Furthermore, COA, FOA, and IGOA converge when the number of iterations is over 20, and FOA has more than 40 when solving SF and GF. In addition, the global optimal solution obtained by COA, FOA, and IGOA is larger than that of FOA.

(e) In Comparison V, for four test functions, the ASR the optimal solution obtained by IWOA is higher than that of other metaheuristic algorithms. This result shows that IWOA is an effective improvement over WOA.

4.2.2. The performance comparison of different operators of IWOA

In this subsection, the performance of two main operational factors (exploration and development) of IWOA is tested. Fig. 5 shows the iterative behavior of IWOA to determine whether it can find the optimal solution of different test functions.

Fig. 5(a), (d), (g), and (j) show the first individual search path of the population based on different test functions for IWOA. For Fig. 5(a), (d), (g), and (j), the first individual of IWOA gradually tends

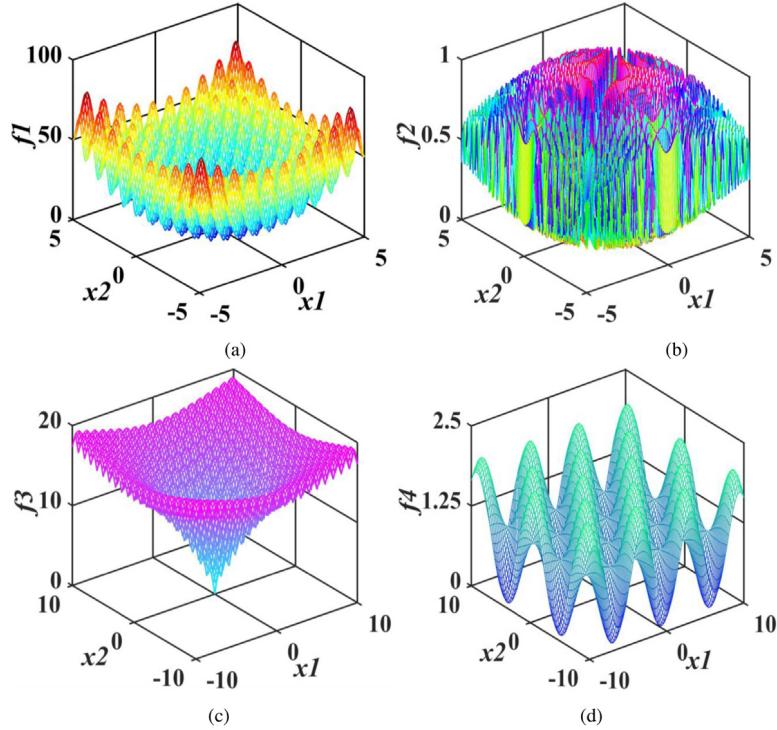


Fig. 3. Two-dimensional graph of four test functions: (a) RF; (b) SF; (c) AF; (d) GF.

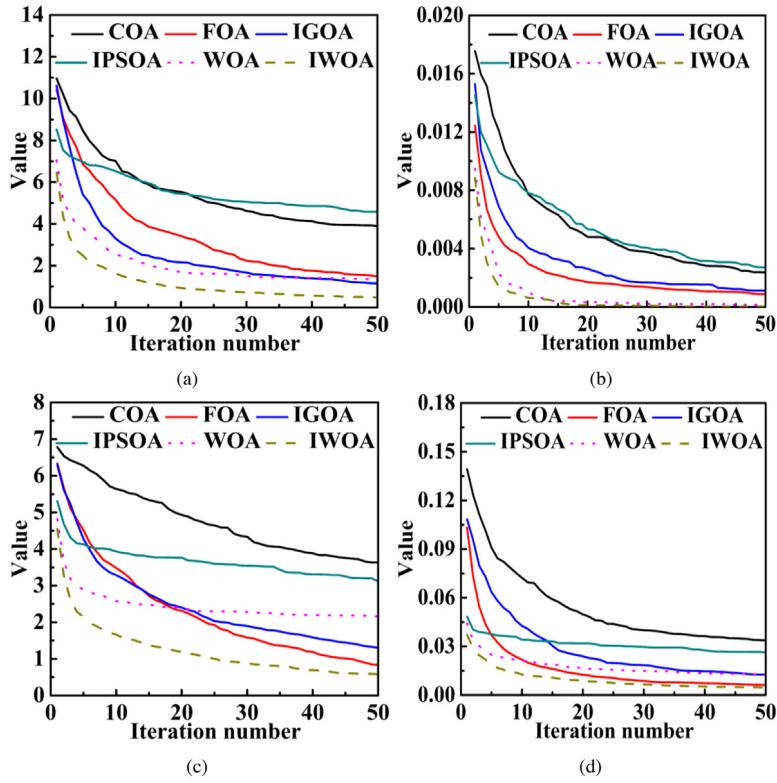


Fig. 4. Comparison of iterative behavior by using IWOA and some metaheuristic algorithms (i.e. COA et al.) based on the four test functions: (a) RF; (b) SF; (c) AF; (d) GF.

to 0 in the search range and the first individual of IWOA has a near fluctuation at 0. Fig. 5(b), (e), (h), and (k) show the search path of the optimal individuals of the population based on different test functions for IWOA. For Fig. 5(b), (e), (h), and (k), we can see that the optimal individuals of IWOA tend to 0 or fluctuate around 0 in the search range, but eventually, all tend to 0. The results of the first-dimensional change

of individuals in the population for different test functions are listed in Fig. 5(c), (f), (i), and (l). It can be seen from Fig. 5(c), (f), (i), and (l) that for different test functions, with the increase of iterations increases, the individual value tends to the solution of the function.

Through the comparison and analysis for Fig. 5, IWOA can gain the global optimal solution for different test functions.

Table 3

Function	Indexes	COA	FOA	IGOA	IPSOA	WOA	IWOA
RF	MMA-1	3.77935	1.51303	1.13302	4.38338	1.48403	0.51382
	MMA-2	10.36743	7.86684	3.91676	6.15157	4.31800	2.01425
	AMA	4.04513	1.90676	1.17080	2.17998	1.18388	0.55109
	STMA	2.23113	1.46183	0.81516	1.32928	0.90282	0.52610
	AG	24.18	33.81	34.04	14.77	15.06	34.58
	ASR	0.03	0.25	0.46	0.03	0.12	0.53
SF	MMA-1	0.00263	0.00076	0.00111	0.00241	0.00026	0.00002
	MMA-2	0.01211	0.00526	0.01318	0.00565	0.00614	0.00313
	AMA	0.00244	0.00058	0.00121	0.00076	0.00071	0.00007
	STMA	0.00297	0.00115	0.00206	0.00124	0.00129	0.00032
	AG	24.51	44.58	24.46	14.7	13.87	47.59
	ASR	0.10	0.45	0.23	0.30	0.24	0.87
AF	MMA-1	3.60384	0.95467	1.46908	3.03626	2.25983	0.52124
	MMA-2	6.92275	3.17573	3.24847	3.59946	5.84841	1.98086
	AMA	3.49479	0.74588	1.31371	2.01516	2.64000	0.13636
	STMA	1.30365	0.72421	0.93582	0.92551	1.20163	0.22844
	AG	24.85	38.22	31.39	17.55	15.48	36.11
	ASR	0.04	0.43	0.25	0.08	0.08	0.96
GF	MMA-1	0.02762	0.00622	0.01032	0.02603	0.01262	0.00318
	MMA-2	0.10250	0.02132	0.03494	0.09392	0.08300	0.02097
	AMA	0.02940	0.00532	0.01076	0.03119	0.02411	0.00306
	STMA	0.01914	0.00464	0.00649	0.01996	0.01635	0.00394
	AG	25.69	42.76	32.62	13.36	16.50	42.88
	ASR	0.04	0.71	0.27	0.06	0.09	0.93

Comparison of iterative behavior for different test functions by using WOA, IWOA, IWOA-1 (exploration factor), and IWOA-2 (exploitation factor) are displayed in Fig. 6. Table 4 shows the performance comparison for different test functions by using WOA, IWOA, IWOA-1, and IWOA-2. By analyzing and comparing Fig. 6 and Table 4, some key results and analysis are shown as follows:

(a) For Fig. 6(a), the convergence bight of IWOA-1 and IWOA-2 is between the convergence curves of WOA and IWOA, and the convergence curve of IWOA is lower than that gained at WOA. Furthermore, the same result appears in Fig. 6(b), (c), and (d). What is more, the convergence curve of IWOA-1 is lower than that of IWOA-2. These results suggest that IWOA is better than WOA, IWOA-1 and IWOA-2, and IWOA-1 is better than IWOA-2.

(b) As can be seen from Table 4, for different test functions, IWOA is lower than WOA, IWOA-1 and IWOA-2 based on MMA-1, MMA-2, AMA, STMA, and ASR, while IWOA is higher than WOA for AG.

(c) As can be seen from Table 4, for RF, IWOA-1 is 0.04040, 0.70925, 0.04046, 0.07243, and 0.08000 lower than IWOA-2 based on MMA-1, MMA-2, AMA, STMA, and AG and IWOA-1 is 0.02 higher than IWOA-2 based on ASR. For SF, IWOA-1 is 0.00001, 0.00022, 0.00001, 0.00002, and 0.10000 lower than IWOA-2 based on MMA-1, MMA-2, AMA, STMA, and AG and IWOA-1 is 0.03 higher than IWOA-2 based on ASR. For AF, IWOA-1 is 0.14135, 0.13461, 0.04094, 0.14573, and 0.07000 lower than IWOA-2 based on MMA-1, MMA-2, AMA, STMA, and AG and IWOA-1 is 0.02 higher than IWOA-2 based on ASR. For GF, IWOA-1 is 0.00022, 0.00258, 0.00019, 0.00093, and 0.38000 lower than IWOA-2 based on MMA-1, MMA-2, AMA, STMA, and AG and IWOA-1 is 0.05 higher than IWOA-2 based on ASR.

Qualitative comparison of mean values of solutions by using WOA, IWOA, IWOA-1, and IWOA-2 based on different test functions is displayed in Fig. 7. Furthermore, this average is the 10 average of 100 experiments. Fig. 8 gives the quantitative comparison of mean values of Solutions by using WOA, IWOA, IWOA-1, and IWOA-2 based on different test functions.

For Fig. 7(a), the average optimal solution gained by IWOA is the approach to the solution of RF at most points. Besides, for Fig. 7(b), (c), and (d), the average optimal solution gained by IWOA is the approach to the optimal solution of SF, AF, and GF at most points.

For Fig. 8(a), the average optimal solution obtained by IWOA is 0.15178, 0.19218, and 0.97021 lower than that of IWOA-1, IWOA-2, and WOA. For Fig. 8(b), the average optimal solution obtained by

Table 4

The performance comparison in view of different test functions solved by WOA, IWOA, IWOA-1, and IWOA-2.

Function	Indexes	WOA	IWOA	IWOA-1	IWOA-2
RF	MMA-1	1.48403	0.51382	0.66560	0.70600
	MMA-2	4.31800	2.01425	2.31963	3.02888
	AMA	1.18388	0.55109	0.66559	0.70605
	STMA	0.90282	0.52610	0.53880	0.61123
	AG	15.06	34.58	34.30	34.38
	ASR	0.12	0.53	0.52	0.50
SF	MMA-1	0.00026	0.00002	0.00011	0.00012
	MMA-2	0.00614	0.00313	0.00468	0.00490
	AMA	0.00071	0.00007	0.00012	0.00013
	STMA	0.00129	0.00032	0.00035	0.00037
	AG	13.87	47.59	47.40	47.50
	ASR	0.24	0.87	0.82	0.79
AF	MMA-1	2.25983	0.52124	0.60132	0.74267
	MMA-2	5.84841	1.98086	4.08465	4.21926
	AMA	2.64000	0.13636	0.61342	0.65436
	STMA	1.20163	0.22844	0.71184	0.85757
	AG	15.48	36.11	36.01	36.08
	ASR	0.38	0.96	0.95	0.93
GF	MMA-1	0.01262	0.00318	0.00566	0.00588
	MMA-2	0.08300	0.02097	0.02715	0.02973
	AMA	0.02411	0.00306	0.00989	0.01008
	STMA	0.01635	0.00394	0.00675	0.00768
	AG	16.50	42.88	42.08	42.46
	ASR	0.35	0.93	0.91	0.86

IWOA is 0.00009, 0.00010, and 0.00024 lower than that of IWOA-1, IWOA-2, and WOA. For Fig. 8(c), the average optimal solution obtained by IWOA is 0.08008, 0.22143, and 1.73859 lower than that of IWOA-1, IWOA-2, and WOA. For Fig. 8(d), the average optimal solution obtained by IWOA is 0.00248, 0.00270, and 0.00944 lower than that of IWOA-1, IWOA-2, and WOA.

4.2.3. Results analysis

In this experiment, two main results are gained. The first is that the performance of IWOA is superior to that of COA, FOA, WOA, IGOA, and IPSOA. The second is that the average optimal solution of IWOA-1 is less than that of IWOA-2 and the optimal solution of IWOA is lower than that of IWOA-1 and IWOA-2.

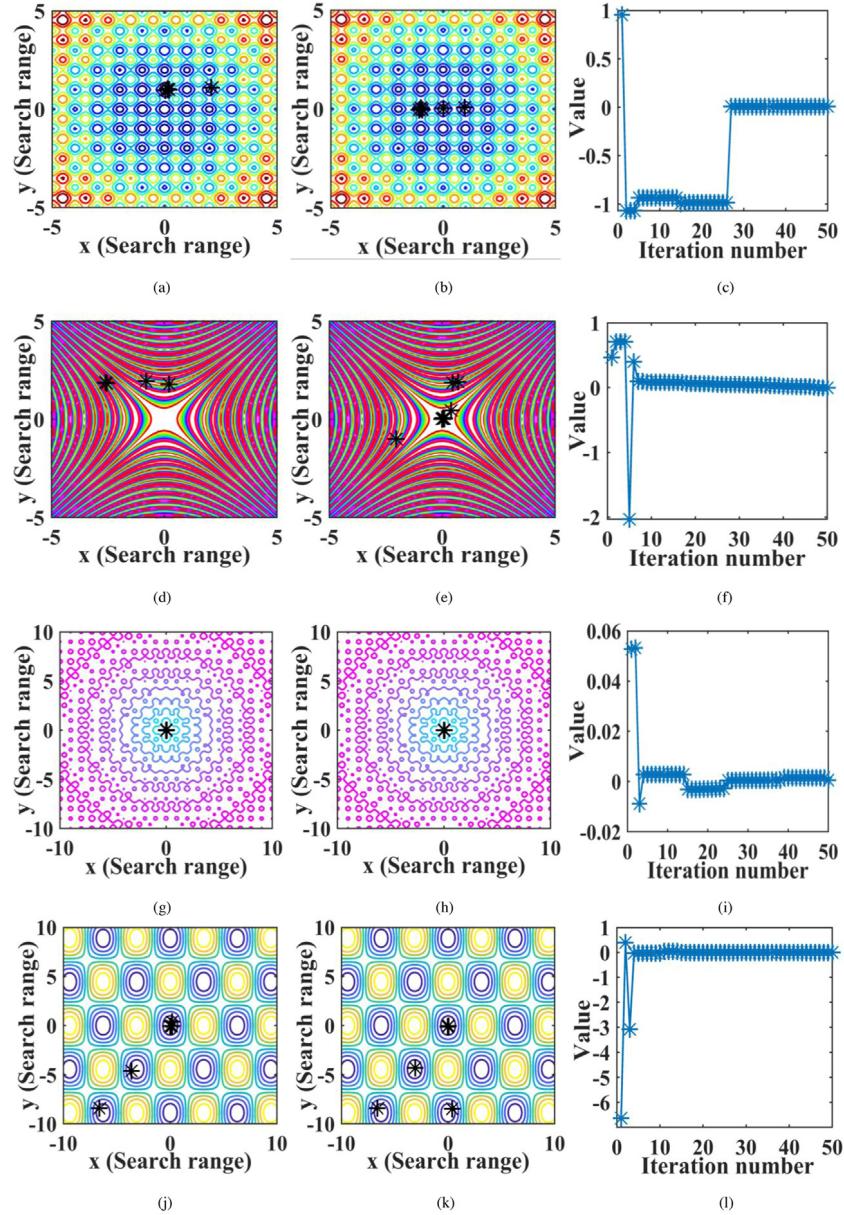


Fig. 5. Iterative behavior based on IWOA for the different function: the search range of the first individual in the population for RF, SF, AF, and GF: (a), (d), (g), and (j); the search range of the best individual in the population for RF, SF, AF, and GF: (b), (e), (h), and (k); the first dimensional change of individuals in population for RF, SF, AF, and GF: (c), (f), (i), and (l).

The reason for the first experimental result is that IWOA is an improvement of WOA, which changes the linear weight to the adaptive weight and increases the jumping behavior. Compared with COA, FOA, IGOA, and IPSOA, WOA increases exploration and exploitation factors.

The second experimental result is due to the different renewal formulas of IWOA-1 and IWOA-2. If $p < P_i$, the renewal equation adopted is $\mathbf{X}^*(t) - AD$, that is, IWOA-1. If $p \geq P_i$, the renewal equation adopted is $\mathbf{X}^*(t) + D_p e^{bt} \cos(2\pi t)$. Different renewal equation adopted will directly lead to different experimental results, that is, the average optimal solution obtained by IWOA-1 is better than that obtained by IWOA-2.

4.3. Experiment II: The performance comparison of different forecasting model

4.3.1. Chaotic characteristic recognition and phase space reconstruction

The maximum Lyapunov exponent is calculated by small data quantity method and the number of evolution steps i is 50, the relationship of distance logarithmic mean $y(i)$ and step number i is listed in Fig. 9(a).

The maximum Lyapunov exponent is 0.0133, which is higher than 0 and it proves that the short-term natural gas consumption time series has obvious chaotic characteristics. Therefore, it can be predicted by the chaotic prediction model.

The $\Delta S(\tau)$ and $S_{\text{our}}(\tau)$ obtained by the C-C method are displayed in Fig. 9(b) and (c). For Fig. 9(b), when τ is 3, $\Delta S(\tau)$ obtains the first minimum, so the optimal delay time is 3. From Fig. 9(c), $S_{\text{our}}(\tau)$ obtains the minimum value and based on the embedded time window equation $\tau_w = (m - 1)\tau$, embedded dimension which is 9 can be obtained.

In summary, the optimal delay time is 3, and the embedded dimension is 9. So, the phase space is reconstructed as below:

$$Y(i) = [x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)] \quad (46)$$

where i is $1, 2, \dots, N$, and $Y(i)$ is reconstructed a sample of phase space, and take m as 9, τ as 3.

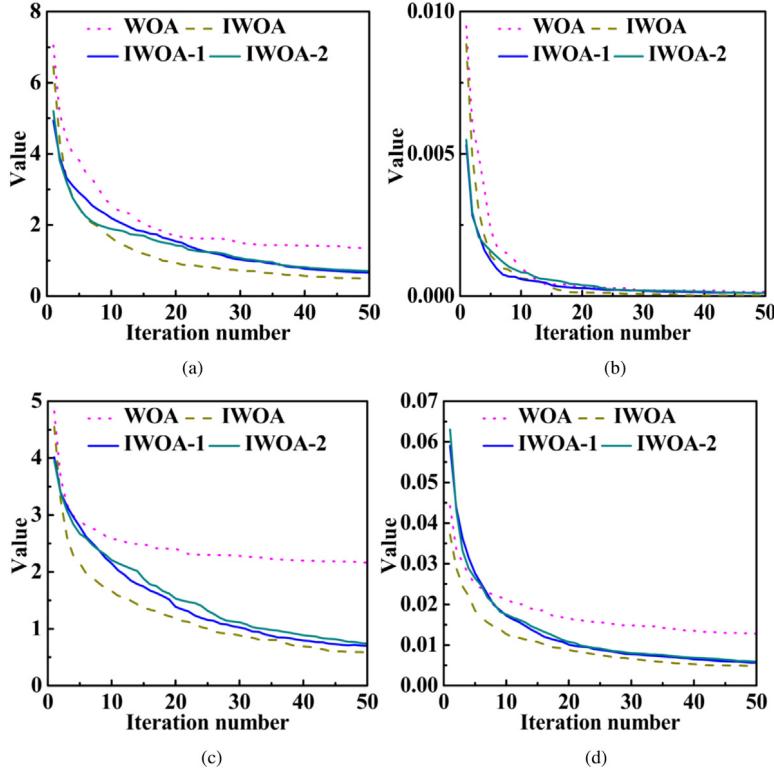


Fig. 6. Comparison of iterative behavior for different test function by using WOA, IWOA, IWOA-1, and IWOA-2: (a) RF; (b) SF; (c) AF; (d) GF.

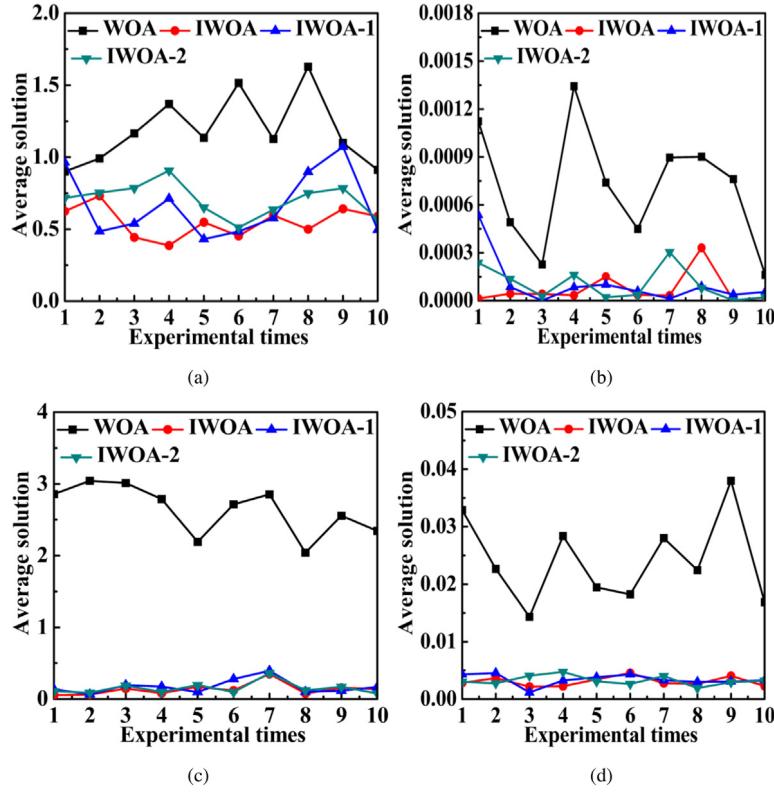


Fig. 7. Qualitative comparison of mean values of solutions of by using WOA, IWOA, IWOA-1, and IWOA-2 based on different test function: (a) RF; (b) SF; (c) AF; (d) GF.

4.3.2. Predictive performance comparison of the six models

In this subsection, IWOA is used to optimize three key parameters (m , τ , the order of Volterra adaptive filter). The three parameters before optimization are set to 3, 9, and 2. Furthermore, the three

optimized parameters are 3, 10, 3. The experimental results for the training sample are depicted in Fig. 10–12, and Tables 5 and 6. Furthermore, Fig. 10 gives the qualitative comparison of results obtained by different forecasting models. Figs. 11 and 12 give the fitting value

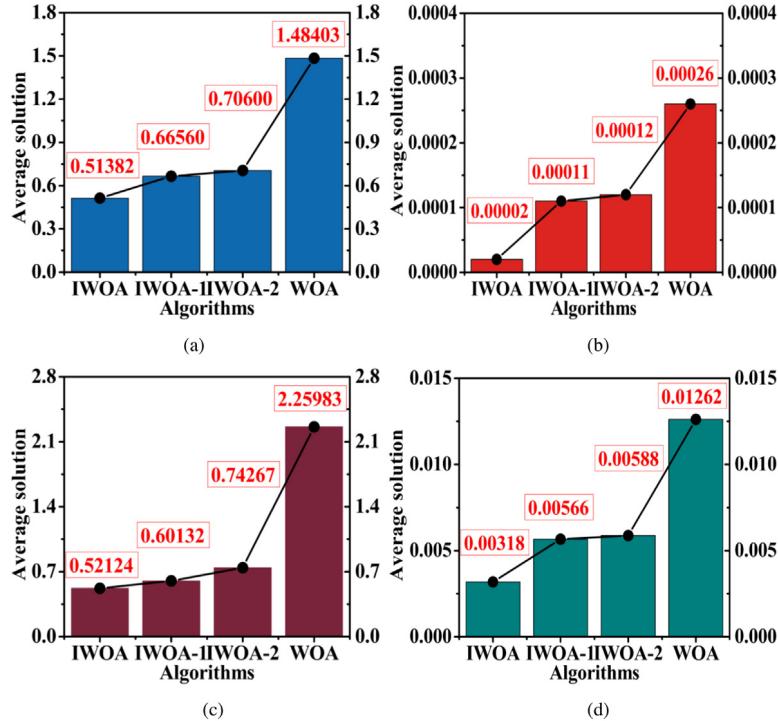


Fig. 8. Quantitative comparison of mean values of Solutions by using WOA, IWOA, IWOA-1, and IWOA-2 based on different test function: (a) RF; (b) SF; (c) AF; (d) GF.

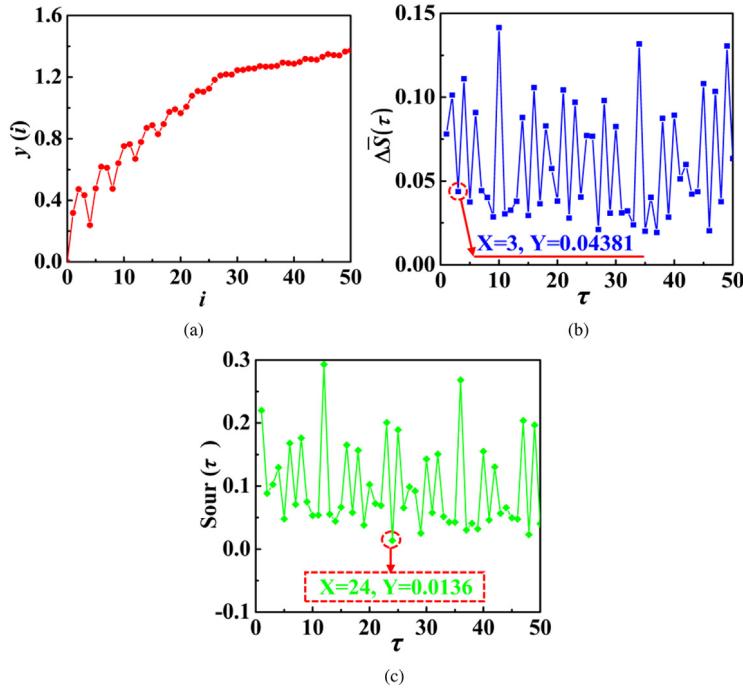


Fig. 9. Calculation results of small data quantity method and C-C method: (a) chaotic character recognition; (b) and (c) phase space reconstruction.

and R^2 obtained by different forecasting models. Tables 5 and 6 show the result comparison by using different forecasting models based on the test sample.

From Tables 5 and 6 as well as Figs. 10, 11, and 12, it can be clearly seen that the developed model is more suitable than the compared models. To deeply analyze the performance of the proposed coupling model, more comparison researches are illustrated as follows.

(a) Comparison I is adopted to qualitatively determine the performance of the developed forecasting model and other prediction models.

From Fig. 10(a), for the training sample, the training values obtained by the developed model and other prediction models can well follow the changing trend of the actual values. In addition, for Fig. 10(b), except for the forecasting values obtained by GS-LSSVM, the predicted values obtained by the other five prediction models can well follow the changing trend of the actual value.

(b) Comparison II is established to the prediction accuracy of the training sample between the proposed and other five models. In Table 5, the range of the three evaluation indexes from low to high is

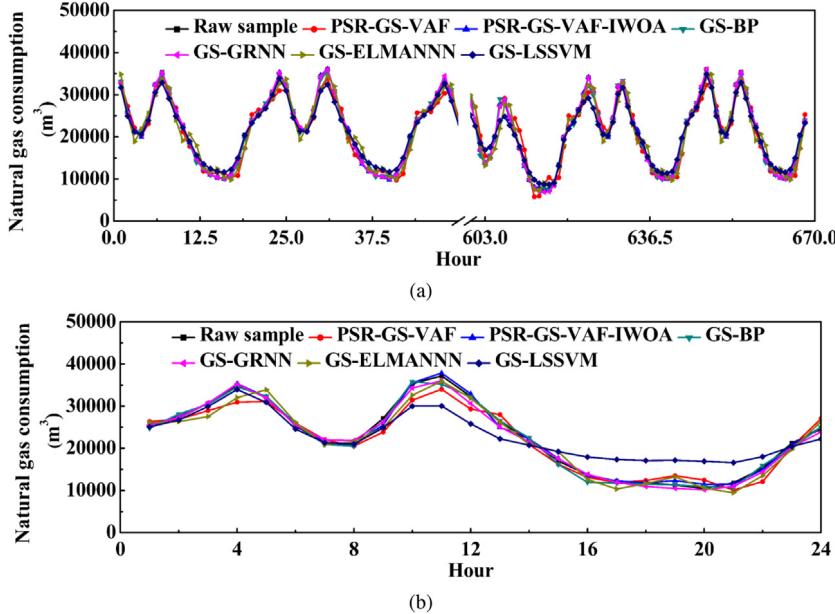


Fig. 10. The qualitative comparison of results obtained by different forecasting models: (a) training sample; (b) test sample.

$238.9658 \rightarrow 2663.9770$, $3.3789 \rightarrow 32.2796$, and $1.3203 \rightarrow 10.6241$ lower than the other five forecasting models. Furthermore, in Table 5, the range of the three evaluation indexes from low to high is $154.0904 \rightarrow 3147.4930$, $3.7625 \rightarrow 33.2578$, and $0.8128 \rightarrow 15.0219$ lower than other five forecasting models. In addition, the R^2 of PSR-GS-VAF-IWOA is higher than that of other forecasting models.

(c) Comparison II is established to the training and test sample between the proposed and other five models. For Table 6, the R^2 of PSR-GS-VAF, PSR-GS-VAF-IWOA, GS-BPNN, GS-GRNN, GS-ELMANN, and GS-LSSVM is 0.9940, 0.9992, 0.9891, 0.9978, 0.9414, and 0.9833 for training sample. And, the R^2 of PSR-GS-VAF-IWOA is 0.0052, 0.0101, 0.0014, 0.0578, and 0.0159 higher than that of PSR-GS-VAF, GS-BPNN, GS-GRNN, GS-ELMANN, and GS-LSSVM. In addition, For Fig. 10, it can be found that the R^2 of PSR-GS-VAF, PSR-GS-VAF-IWOA, GS-BPNN, GS-GRNN, GS-ELMANN, and GS-LSSVM is 0.9494, 0.9946, 0.9919, 0.9939, 0.9631, and 0.9006 for the test sample. And, the R^2 of PSR-GS-VAF-IWOA is 0.0452, 0.0027, 0.0007, 0.0315, and 0.0940 higher than that of PSR-GS-VAF, GS-BPNN, GS-GRNN, GS-ELMANN, and GS-LSSVM.

4.3.3. Results analysis

In this experiment, two main experimental results are obtained. The first is that the short-term natural gas consumption time series has chaotic characteristics. The second is that the prediction accuracy of PSR-GS-VAF-IWOA is higher than that of PSR-GS-VAF, GS-BPNN, GS-GRNN, GS-ELMANN, and GS-LSSVM.

In recent years, people pay more and more attention to the analysis of nonlinear systems. The study of chaos from time series began with the reconstruction phase space theory proposed (Packard et al., 1980). Therefore, in this experiment, the chaotic characteristics of the short-term natural gas consumption time series are analyzed, and the chaotic characteristics are determined.

The second result is that for PSR-GS-VAF, the parameters of this model are fixed, but PSR-GS-VAF-IWOA uses IWOA to optimize the three main parameters and determine the optimal parameters.

5. Discussions

In this section, the discussion is carried out, which includes two subsections: complexity analysis based on Experiment I and different forecasting models testing based on Experiment II.

5.1. Complexity analysis based on Experiment I

In this subsection, the computing complexity of the algorithm involved in Experiment I is discussed. The trend of operands with the experimental times, maximum number of iterations, and population number and the trend of the worst operands with experimental times are based on different metaheuristic algorithm is displayed in Fig. 13. In addition, the average execution time of different metaheuristic algorithms is depicted in Fig. 14.

In the analysis of computing complexity, the maximum number of iterations is expressed in m , the population number is expressed in n , and the experimental times are expressed in t . In this subsection, the big O method is adopted to analyze the computational complexity of different algorithms in view of Experiment I. Through analysis, the complexity of COA, FOA, IGOA, IPSOA, WOA, IWOA, IWOA-1, and IWOA-2 is $O(13t + 7tn + 11tnm)$, $O(13t + 12tn + 25tnm)$, $O(10t + 14tn + 26tnm)$, $O(13t + 12tn + 12tnm)$, $O(14t + 14tn + 22tnm)$, $O(16t + 16tn + 24tnm)$, $O(15t + 16tn + 28tnm)$, $O(15t + 18tn + 30tnm)$. For Fig. 13(a), as the number of experimental times increases, the operands of different metaheuristic algorithms increases, and the order of their changes from low to high is COA, IPSOA, WOA, FOA, IGOA, IWOA-1, IWOA-2, and IWOA. For Fig. 13(b), as the number of the maximum iterations raises, the operands of different metaheuristic algorithms increases, and the order of their changes from low to high is COA, IPSOA, WOA, FOA, IGOA, IWOA-1, IWOA-2, and IWOA. For Fig. 13(c), as the number of populations, the operands of different metaheuristic algorithms increases, and the order of their changes from low to high is COA, IPSOA, WOA, FOA, IGOA, IWOA-1, IWOA-2, and IWOA.

Because of $n < t$ and $m < t$, so the worst operands of COA, FOA, IGOA, IPSOA, WOA, IWOA, IWOA-1, and IWOA-2 can be expressed as $O(11t^3)$, $O(25t^3)$, $O(26t^3)$, $O(12t^3)$, $O(22t^3)$, $O(24t^3)$, $O(28t^3)$, $O(30t^3)$. By analyzing Fig. 13(d), the order of the worst operands from low to high is COA, IPSOA, WOA, FOA, IGOA, IWOA-1, IWOA-2, and IWOA.

Based on Fig. 14, it can be clearly obtained that the average execution time of COA to solve the four test functions is the smallest, that of IWOA to solve the four test functions is the largest, and higher than IWOA-1 and IWOA-2, and that of IWOA-2 is higher than that of IWOA-1. The order of average execution time of four test functions solved by different metaheuristic algorithms is COA, IPSOA, WOA, FOA, IGOA, IWOA-1, IWOA-2, and IWOA. This result further validates the complexity sorting of different metaheuristic algorithms using the big O analysis method.

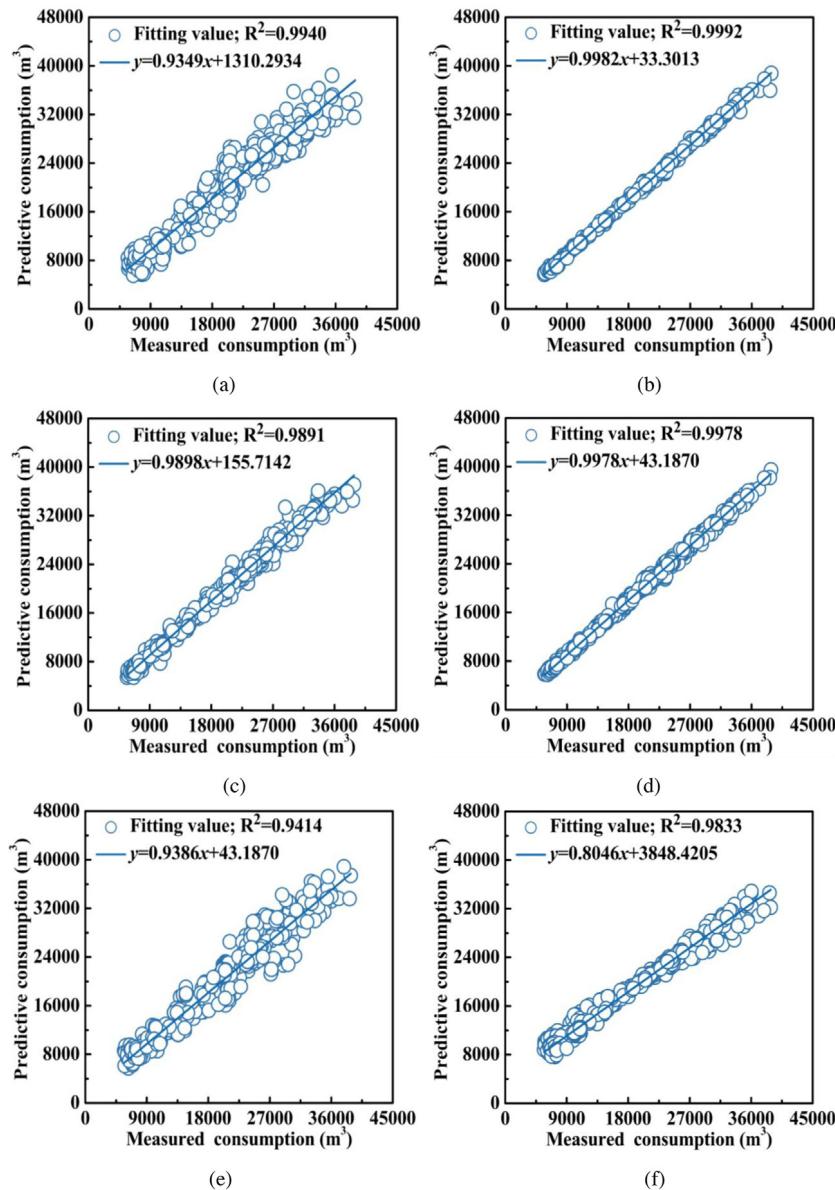


Fig. 11. Fitting value and R^2 obtained by different forecasting model for the training sample: (a) PSR-GS-VAF; (b) PSR-GS-VAF-IWOA; (c) GS-BPNN; (d) GS-GRNN; (e) GS-ELMANN; (f) GS-LSSVM.

Table 5

Result comparison by using different forecasting models based on the training sample.

Indices	PSR-GS-VAF	PSR-GS-VAF-IWOA	GS-BPNN	GS-GRNN	GS-ELMANN	GS-LSSVM
RSME (m^3)	3048.1429	384.1662	813.9551	623.1320	1917.4416	1818.2430
MAE (m^3)	48.9989	16.7193	23.8094	20.0982	39.4486	36.9926
MAPE (%)	12.2403	1.6162	3.1867	2.9365	9.1726	9.6903

Note The values in bold show minimum values of RMSE, MAE, MAPE and R^2 for the training sample.

Table 6

Result comparison by using different forecasting models based on the test sample.

Indices	PSR-GS-VAF	PSR-GS-VAF-IWOA	GS-BPNN	GS-GRNN	GS-ELMANN	GS-LSSVM
RSME (m^3)	3281.9945	599.9940	899.9082	754.0844	1994.7285	3747.4870
MAE (m^3)	51.9916	21.3546	25.8249	25.1171	41.3947	54.6124
MAPE (%)	12.4312	2.3416	3.6413	3.1544	10.8953	17.3635

Note The values in bold show minimum values among RMSE, MAE, MAPE and R^2 for the test sample.

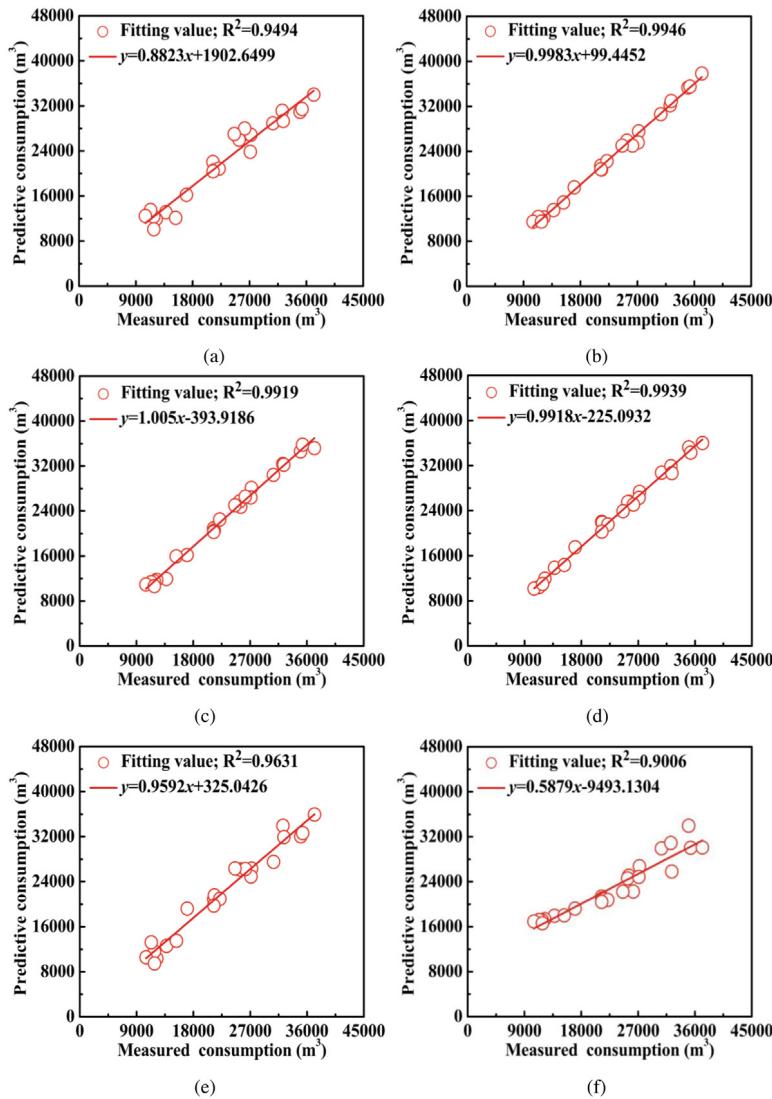


Fig. 12. Fitting value and R^2 obtained by different forecasting model for the test sample: (a) PSR-GS-VAF; (b) PSR-GS-VAF-IWOA; (c) GS-BPNN; (d) GS-GRNN; (e) GS-ELMANN; (f) GS-LSSVM.

5.2. Different forecasting models testing based on Experiment II

Table 7 shows the prediction accuracy levels. According to the results of **Tables 8** and **9**, some analyses and comparisons about the performances of the different forecast models can be discussed in detail as follows:

(b) From **Table 8**, the Q value of the developed model is smaller than that of the other five benchmark forecast models for the training set. According to **Table 8**, the developed model has two accuracy classes for the training set, which is bigger than the accuracy class of the other five benchmark forecast models for training set. In addition, the C value of the developed model is smaller than that of the other five benchmark forecast models for training set. According to **Table 8**, the developed model for training set has one accuracy class, which is bigger than the accuracy class of other five benchmark forecast models for the training set. Furthermore, the p value of the developed model is smaller than that of the other five benchmark forecast models for the training set. According to **Table 8**, the developed model for training set has one accuracy class, which is bigger than the accuracy class of other five benchmark forecast models for training set.

(c) From **Table 9**, the Q value of the developed model is smaller than that of the other models for the test set. According to **Table 9**, the developed model has one accuracy class for the test set, which is bigger

Table 7
Accuracy class among Q , C , and p .

Accuracy class	One	Two	Three	Four
Q	0.01	0.05	0.10	0.20
C	$C \leq 0.35$	$0.35 < C \leq 0.50$	$0.50 < C \leq 0.65$	$0.65 < C$
p	$p \geq 0.95$	$0.95 > p \geq 0.80$	$0.80 > p \geq 0.70$	$0.70 > p$

than the accuracy class of the other five benchmark forecast models for the test set. What is more, the C value of the developed model for the test set is smaller than that of the other models for test set. According to **Table 9**, the developed model for test set has one accuracy class, which is bigger than the accuracy class of other five benchmark forecast models for test set. Besides, the p value of the developed model is smaller than that of the other models for test set. According to **Table 9**, the developed model for test set has one accuracy class, which is bigger than the accuracy class of other five benchmark forecast models for test set.

5.3. Insufficient research and future research

In this study, a novel hybrid prediction model has been proposed, in which normalization, Gaussian smoothing, chaotic characteristics

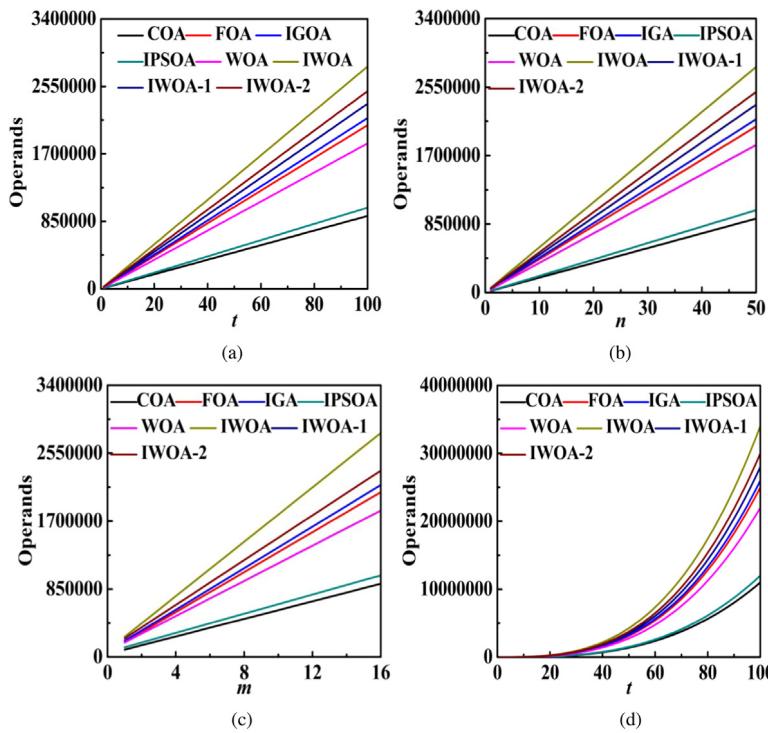


Fig. 13. The trend of operands with the experimental times, maximum number of iterations, and population number and the trend of the worst operands with experimental times are based on different metaheuristic algorithm: (a) the maximum number of iterations; (b) population number; (c) experimental times. (d) the trend of the worst operands with experimental times.

Table 8
Different models checking for the training sample.

Indices	PSR-GS-VAF	PSR-GS-VAF-IWOA	GS-BPNN	GS-GRNN	GS-ELMANN	GS-LSSVM
Q	0.126	0.023	0.039	0.033	0.083	0.144
C	0.479	0.072	0.099	0.077	0.238	0.735
p	0.875	1.000	1.000	1.000	1.000	0.583

Table 9
Different models checking for the test sample.

Indices	PSR-GS-VAF	PSR-GS-VAF-IWOA	GS-BPNN	GS-GRNN	GS-ELMANN	GS-LSSVM
Q	0.123	0.006	0.032	0.016	0.090	0.083
C	0.399	0.028	0.099	0.047	0.237	0.273
p	0.901	1.000	1.000	1.000	0.999	0.978

recognition, phase space reconstruction, Volterra adaptive filter, and IWOA have been applied. However, the research is still insufficient. For example, this novel model should require more test samples. Besides, the efficiency of the developed model needs to be improved.

In the future research, the proposed innovation coupling forecasting model should be applied to more field parameter optimization, such as the parameter optimization of the large molded rock salt specimens (Liu et al., 2019a,b), the structural design optimization of gas wells (Qiao and Wang, 2018), the structural optimization of the rectifying plate system (Peng et al., 2019; Liu et al., 2019a,b; Su et al., 2019), the experimental parameters optimization of 3PE coatings (Qiao et al., 2019d), the parameter optimization of the deep learning algorithm (Qiao et al., 2019e), and, etc.

6. Conclusion

Short term natural gas consumption is a very important basic index for the design of natural gas pipeline network, the preparation of special natural gas planning, and the intelligent scheduling of natural

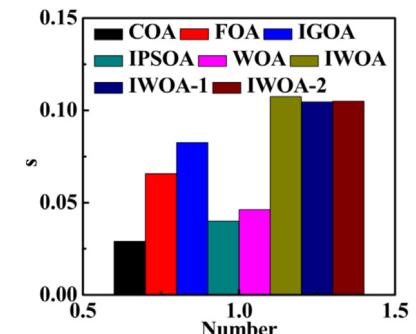


Fig. 14. The average execution time of different metaheuristic algorithms for four test functions.

gas companies. In this study, a novel forecasting model is developed and applied to an example. Some important conclusions are drawn as follows:

(a) The performance of IWOA is superior to that of the cuckoo optimization algorithm etc, exploration and exploitation factor of IWOA based on the different evaluating indicators. However, the average execution time of the IWOA is exceeding that of other metaheuristic algorithms considered for comparison. Furthermore, for exploration and exploitation factor of IWOA, the performance of exploration factor is better than that of exploitation factor.

(b) The performance of the developed hybrid forecast model in view of phase space reconstruction, Gaussian smoothing, Volterra adaptive filter, and IWOA is better than that of Gaussian smoothing-back propagation neural network, Gaussian smoothing-general regression neural network, Gaussian smoothing-Elman neural network, Gaussian smoothing-phase space reconstruction-Volterra adaptive filter, Gaussian smoothing-least squares support vector machine based on different evaluating indicator.

All in all, the prediction model developed has high prediction accuracy in forecasting short-term natural gas consumption but has low efficiency. Therefore, improving the efficiency of this prediction model, testing the model more, and applying IWOA to parameter optimization in other fields are the next main research direction.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to <https://doi.org/10.1016/j.engappai.2019.103323>.

Acknowledgments

This study is funded the high-level talents start-up project of north China university of water resources and electric power with the grand number of 40691. Authors would like to thank the editor and the anonymous referees for their valuable comments and criticisms.

References

- Abdel-Basset, M., Manogaran, G., El-Shahat, D., Mirjalili, S., 2018. A hybrid whale optimization algorithm based on local search strategy for the permutation flow shop scheduling problem. Future Gener. Comput. Syst. 85, 129–145.
- Akpinar, M., Adak, M.F., Yumusak, N., 2016. Forecasting natural gas consumption with hybrid neural networks—Artificial bee colony. In: 2016 2nd International Conference on Intelligent Energy and Power Systems (IEPS). IEEE, pp. 1–6.
- Aljarrah, I., Faris, H., Mirjalili, S., 2018. Optimizing connection weights in neural networks using the whale optimization algorithm. Soft Comput. 22 (1), 1–15.
- Augusteijn, H.E., van Aert, R., van Assen, M.A., 2019. The effect of publication bias on the q test and assessment of heterogeneity. Psychol. Methods 24 (1), 116.
- Azadeh, A., Asadzadeh, S.M., Ghanbari, A., 2010. An adaptive network-based fuzzy inference system for short-term natural gas demand estimation: uncertain and complex environments. Energy Policy 38 (3), 1529–1536.
- Azari, A., Sharaty-Niassar, M., Alborzi, M., 2012. Short-term and medium-term gas demand load forecasting by neural networks. Iranian J. Chem. Chem. Eng. (IJCCE) 31 (4), 77–84.
- Bai, Y., Li, C., 2016. Daily natural gas consumption forecasting based on a structure-calibrated support vector regression approach. Energy Build. 127, 571–579.
- Bai, Y., Li, Y., Zeng, B., Li, C., Zhang, J., 2019. Hourly PM_{2.5} concentration forecast using stacked autoencoder model with emphasis on seasonality. J. Cleaner Prod. 224, 739–750.
- Brabec, M., Konár, O., Pelikán, E., Malý, M., 2008. A nonlinear mixed effects model for the prediction of natural gas consumption by individual customers. Int. J. Forecast. 24 (4), 659–678.
- Cao, L., 1997. Practical method for determining the minimum embedding dimension of a scalar time series. Physica D 110 (1–2), 43–50.
- Dai, Y., Guo, J., Yang, L., You, W., 2019. A new approach of intelligent physical health evaluation based on GRNN and BPNN by using a wearable smart bracelet system. Procedia Comput. Sci. 147, 519–527.
- Deetman, S., Hof, A.F., Pfluger, B., van Vuuren, D.P., Girod, B., van Ruijven, B.J., 2013. Deep greenhouse gas emission reductions in Europe: Exploring different options. Energy Policy 55, 152–164.
- Demirel, Ö.F., Zaim, S., Çalışkan, A., Özuyar, P., 2012. Forecasting natural gas consumption in Istanbul using neural networks and multivariate time series methods. Turkish J. Electr. Eng. Comput. Sci. 20 (5), 695–711.
- Du, P., Wang, J., Guo, Z., Yang, W., 2017. Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting. Energy Convers. Manage. 150, 90–107.
- Du, P., Wang, J., Yang, W., Niu, T., 2019. A novel hybrid model for short-term wind power forecasting. Appl. Soft Comput. J. 2019, <http://dx.doi.org/10.1016/j.asoc.2019.03.035>.
- Ervural, B.C., Beyca, O.F., Zaim, S., 2016. Model estimation of ARMA using genetic algorithms: A case study of forecasting natural gas consumption. Proc.-Soc. Behav. Sci. 235, 537–545.
- Guo, T., He, W., Jiang, Z., Chu, X., Malekian, R., Li, Z., 2019. An improved LSSVM model for intelligent prediction of the daily water level. Energies 12 (1), 112.
- Hippert, H.S., Pedreira, C.E., Souza, R.C., 2001. Neural networks for short-term load forecasting: A review and evaluation. IEEE Trans. Power Syst. 16 (1), 44–55.
- Ivezic, D., 2006. Short-term natural gas consumption forecast. FME Trans. 34 (3), 165–169.
- Karimi, H., Dastranj, J., 2014. Artificial neural network-based genetic algorithm to predict natural gas consumption. Energy Syst. 5 (3), 571–581.
- Kashani, M.H., Ghorbani, M.A., Dinparshoh, Y., Shahmorad, S., 2016. Integration of Volterra model with artificial neural networks for rainfall-runoff simulation in forested catchment of northern Iran. J. Hydrol. 540, 340–354.
- Khotanzad, A., Elragal, H., 1999. Natural gas load forecasting with combination of adaptive neural networks. In: International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339). In: IJCNN'99, vol. 6, IEEE, pp. 4069–4072.
- Khotanzad, A., Elragal, H., Lu, T.L., 2000. Combination of artificial neural-network forecasters for prediction of natural gas consumption. IEEE Trans. Neural Netw. 11 (2), 464–473.
- Kizilaslan, R., Karlik, B., 2008. Comparison neural networks models for short term forecasting of natural gas consumption in Istanbul. In: 2008 First International Conference on the Applications of Digital Information and Web Technologies (ICADIWT). IEEE, pp. 448–453.
- Liu, E., Li, W., Cai, H., Peng, S., 2019a. Formation mechanism of trailing oil in product oil pipeline. Processes 7 (1), 7.
- Liu, Z., Li, W., Sun, W., 2013. A novel method of short-term load forecasting based on multiwavelet transform and multiple neural networks. Neural Comput. Appl. 22 (2), 271–277.
- Liu, W., Zhang, Z., Chen, J., Fan, J., Jiang, D., Jjk, D., Li, Y., 2019b. Physical simulation of construction and control of two butted-well horizontal cavern energy storage using large molded rock salt specimens. Energy 185, 682–694.
- Lu, K.H., Hong, C.M., Xu, Q., 2019. Recurrent wavelet-based Elman neural network with modified gravitational search algorithm control for integrated offshore wind and wave power generation systems. Energy 170, 40–52.
- Mafarja, M.M., Mirjalili, S., 2017. Hybrid whale optimization algorithm with simulated annealing for feature selection. Neurocomputing 260, 302–312.
- Mafarja, M., Mirjalili, S., 2018. Whale optimization approaches for wrapper feature selection. Appl. Soft Comput. 62, 441–453.
- Malakoti-Moghadam, M., Askarzadeh, A., Rashidinejad, M., 2019. Transmission and generation expansion planning of energy hub by an improved genetic algorithm. Energy Sources A 1–15.
- Mehne, H.H., Mirjalili, S., 2018. A parallel numerical method for solving optimal control problems based on whale optimization algorithm. Knowl.-Based Syst. 151, 114–123.
- Merkel, G., Povinelli, R., Brown, R., 2018. Short-term load forecasting of natural gas with deep neural network regression. Energies 11 (8), 2008.
- Mirjalili, S., Lewis, A., 2016. The whale optimization algorithm. Adv. Eng. Softw. 95, 51–67.
- Mousavi, S.M., Alikar, N., Tavana, M., Di Caprio, D., 2019. An improved particle swarm optimization model for solving homogeneous discounted series-parallel redundancy allocation problems. J. Intell. Manuf. 30 (3), 1175–1194.
- Packard, N.H., Crutchfield, J.P., Farmer, J.D., Shaw, R.S., 1980. Controlling chaos. Phys. Rev. Lett. 45, 12–18.
- Panapakidis, I.P., Dagoumas, A.S., 2017. Day-ahead natural gas demand forecasting based on the combination of wavelet transform and ANFIS/genetic algorithm/neural network model. Energy 118, 231–245.
- Peng, S., Chen, Q., Zheng, C., Liu, E., 2019. Analysis of particle deposition in a new-type rectifying plate system during shale gas extraction. Energy Sci. Eng. 2019, <http://dx.doi.org/10.1002/ese3.543>.
- Potočnik, P., Soldo, B., Šimunović, G., Šarić, T., Jeromen, A., Govekar, E., 2014. Comparison of static and adaptive models for short-term residential natural gas forecasting in Croatia. Appl. Energy 129, 94–103.
- Qiao, W., Bingfan, L., Zhangyang, K., 2019d. Differential scanning calorimetry and electrochemical tests for the analysis of delamination of 3PE coatings. Int. J. Electrochem. Sci. 14, 7389–7400. <http://dx.doi.org/10.20964/2019.08.05>.
- Qiao, W., Huang, K., Azimi, M., Han, S., 2019a. A novel hybrid prediction model for hourly gas consumption in supply side based on improved whale optimization algorithm and relevance vector machine. IEEE Access 7, 88218–88230. <http://dx.doi.org/10.1109/ACCESS.2019.2918156>.
- Qiao, W., Lu, H., Zhou, G., Azimi, M., Yang, Q., Tian, W., 2020. A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer. J. Cleaner Prod. 244, 118612. <http://dx.doi.org/10.1016/j.jclepro.2019.118612>.
- Qiao, W., Tian, W., Tian, Y., Yang, Q., Wang, Y., Zhang, J., 2019e. The forecasting of PM 2.5 using a hybrid model based on wavelet transform and an improved deep learning algorithm. IEEE Access 7, 142814–142825. <http://dx.doi.org/10.1109/ACCESS.2019.2944755>.
- Qiao, W., Wang, H., 2018. Analysis of the wellhead growth in HPHT gas wells considering the multiple annuli pressure during production. J. Nat. Gas Sci. Eng. 50, 43–54. <http://dx.doi.org/10.1016/j.jngse.2017.10.028>.
- Qiao, W., Yang, Z., 2019b. Modified dolphin swarm algorithm based on chaotic maps for solving high-dimensional function optimization problems. IEEE Access 7, 110472–110486. <http://dx.doi.org/10.1109/ACCESS.2019.2931910>.
- Qiao, W., Yang, Z., 2019c. Solving large-scale function optimization problem by using a new metaheuristic algorithm based on quantum dolphin swarm algorithm. IEEE Access 7, 138972–138989. <http://dx.doi.org/10.1109/ACCESS.2019.2942169>.
- Rosenstein, M.T., Collins, J.J., De Luca, C.J., 1993. A practical method for calculating largest Lyapunov exponents from small data sets. Physica D 65 (1–2), 117–134.
- Sabo, K., Scitovski, R., Vazler, I., Zekić-Sušić, M., 2011. Mathematical models of natural gas consumption. Energy Convers. Manage. 52 (3), 1721–1727.
- Shi, W.Y., 2013. Research on gas load forecasting using artificial neural network. In: Advanced Materials Research, Vol. 717. Trans Tech Publications, pp. 423–427.
- Soldo, B., 2012. Forecasting natural gas consumption. Appl. Energy 92, 26–37.

- Soldo, B., Potočnik, P., Šimunović, G., Šarić, T., Govekar, E., 2014. Improving the residential natural gas consumption forecasting models by using solar radiation. *Energy Build.* 69, 498–506.
- Su, Z., Liu, E., Xu, Y., Xie, P., Shang, C., Zhu, Q., 2019. Flow field and noise characteristics of manifold in natural gas transportation station. *Oil Gas Sci. Technol.–Rev. IFP Energ. Nouv.* 74 (70).
- Suganthi, L., Samuel, A.A., 2012. Energy models for demand forecasting—A review. *Renew. Sustain. Energy Rev.* 16 (2), 1223–1240.
- Szoplik, J., 2015. Forecasting of natural gas consumption with artificial neural networks. *Energy* 85, 208–220.
- Taşpinar, F., Celebi, N., Tutkun, N., 2013. Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods. *Energy Build.* 56, 23–31.
- Tronarp, F., Särkkä, S., 2019. Iterative statistical linear regression for Gaussian smoothing in continuous-time non-linear stochastic dynamic systems. *Signal Process.* 159, 1–12.
- Viet, N.H., Mandziuk, J., 2003. Neural and fuzzy neural networks for natural gas consumption prediction. In: 2003 IEEE XIII Workshop on Neural Networks for Signal Processing (IEEE Cat. No. 03TH8718). IEEE, pp. 759–768.
- Wang, J., Yang, W., Du, P., Li, Y., 2018. Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system. *Energy* 148, 59–78.
- Wei, N., Li, C., Duan, J., Liu, J., Zeng, F., 2019a. Daily natural gas load forecasting based on a hybrid deep learning model. *Energies* 12 (2), 218.
- Wei, N., Li, C., Li, C., Xie, H., Du, Z., Zhang, Q., Zeng, F., 2019b. Short-term forecasting of natural gas consumption using factor selection algorithm and optimized support vector regression. *J. Energy Resour. Technol.* 141 (3), 032701.
- Wei, N., Li, C., Peng, X., Li, Y., Zeng, F., 2019c. Daily natural gas consumption forecasting via the application of a novel hybrid model. *Appl. Energy* 250, 358–368.
- Wei, N., Li, C., Peng, X., Zeng, F., Lu, X., 2019d. Conventional models and artificial intelligence-based models for energy consumption forecasting: A review. *J. Pet. Sci. Eng.* 106187.
- Wolf, A., Swift, J.B., Swinney, H.L., Vastano, J.A., 1985. Determining Lyapunov exponents from a time series. *Physica D* 16 (3), 285–317.
- Yu, F., Wang, Z.Q., Xu, X.Z., 2014. Short-term gas load forecasting based on wavelet BP neural network optimized by genetic algorithm. In: *Applied Mechanics and Materials*, Vol. 631. Trans Tech Publications, pp. 79–85.
- Yu, F., Xu, X., 2014. A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. *Appl. Energy* 134, 102–113.
- Zeng, Z., Bao, H., Wen, Z., Zhu, W., 2019. Object tracking using the particle filter optimized by the improved artificial fish swarm algorithm. *Int. J. Intell. Inf. Database Syst.* 12 (1–2), 6–19.
- Zhang, C., Liu, Y., Zhang, H., Huang, H., Zhu, W., 2010. Research on short-term gas load forecasting based on support vector machine model. In: *Life System Modeling and Intelligent Computing*. Springer, Berlin, Heidelberg, pp. 390–399.
- Zhou, H., Su, G., Li, G., 2011. Forecasting daily gas load with OIHF-Elman neural network. *Procedia Comput. Sci.* 5, 754–758.
- Zhu, L., Li, M.S., Wu, Q.H., Jiang, L., 2015. Short-term natural gas demand prediction based on support vector regression with false neighbours filtered. *Energy* 80, 428–436.
- Zhu, X., Wang, N., 2019. Cuckoo search algorithm with onlooker bee search for modeling PEMFCs using T2FNN. *Eng. Appl. Artif. Intell.* 85, 740–753.