

Energy Saving Diagnosis Model of Petrochemical Plant Based on Intelligent Curvelet Support Vector Machine

bin zhao (✉ zbzbz0203288@163.com)

Liaoning University of Petroleum and Chemical Technology

dou qin

Liaoning Shihua University: Liaoning University of Petroleum and Chemical Technology

Diankui Gao

Liaoning Shihua University: Liaoning University of Petroleum and Chemical Technology

lizhi xu

Liaoning Shihua University: Liaoning University of Petroleum and Chemical Technology

Research Article

Keywords: energy saving diagnosis, petrochemical plant, twin Curvelet support vector machine

Posted Date: April 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-412350/v1>

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

[Read Full License](#)

Energy Saving Diagnosis Model of Petrochemical Plant Based on Intelligent Curvelet Support Vector Machine

Bin Zhao*, Dou Qin, Diankui Gao, Lizhi Xu

(School of Mechanical Engineering, Liaoning Shihua University, Fushun, Liaoning, 113001, China)

zbzbz0203288@163.com

Abstract: The energy and resources saving has become a major task of petrochemical enterprises, it is necessary to construct the energy saving diagnostic system for understand the real time operation information of petrochemical plant and provide theoretical basis for taking energy saving measures. The energy saving diagnosis process of petrochemical plant based on twin Curvelet support vector machine optimized by hybrid glowworm swarm algorithm is designed. The Curvelet kernel function is constructed based on curvelet transform to establish theory model of twin curvelet support vector machine. In order to improve the prediction precision of the twin curvelet support vector machine, the hybrid glowworm swarm optimization algorithm is constructed based on simulated annealing simulation to optimize the parameters of the twin Curvelet support vector machine. Finally, a petrochemical plant is used as research object to carry out diagnosis simulation analysis, and results showed that the proposed prediction model can effectively improve diagnostic effectiveness of the energy saving effect of petrochemical plant.

Keywords: energy saving diagnosis; petrochemical plant; twin Curvelet support vector machine

1 Introduction

With development of economy, energy crisis and environmental pollution are rising, which have become two difficult problems facing the world. In order to cope with growing energy demand and environmental pressure, China has taken the optimization and adjustment of the industry and economic structure, and brought in a number of measures. The petrochemical enterprises consume a lot of energy source, the energy consumption can directly affect competitiveness of petrochemical enterprises, therefore energy and resources saving has become a major task of petrochemical enterprises.

On energy-saving, sorts of energy-saving technologies and devices are applied in petrochemical plant, but often overlook the energy saving potential, therefore the energy saving and efficiency improvement are short of solution, and the energy consumption is uneven. The energy saving and efficiency improvement of petrochemical plant. On condition that the operating conditions of the petrochemical plant are determined, the manager of petrochemical enterprises can grasp and identify the potential problems and formulate the energy innovation projects, and then the energy saving and efficiency improvement can be achieved to ensure the safe and efficient operation of petrochemical plant. It is necessary to establish a scientific and rational diagnostic system of energy-saving of petrochemical plant. The energy saving diagnosis can understand the real time operation information of petrochemical plant through analyzing the data for energy saving and consumption reducing of petrochemical enterprises. The weakness of energy utilization of petrochemical enterprises can be cleared, the existing problems in operation and management of petrochemical plant can be identified according to energy saving diagnosis results, and then the

energy-saving transformation measures can be made, which can achieve the efficient use of resources. The energy diagnosis can provide the basis for management decision-making of petrochemical enterprises.

In recent years, the energy saving evaluation of various industries has been concerned by some scientists, many outstanding achievements have come out. Shi Wantao et al. proposed a novel evaluation method based on exergy analysis, the effectiveness of the proposed method is verified through comparison and analysis [1]. Kasturi Srikant and Kannappan Bakthavatsalam Ayalar conducted energy audits, and the evaluated energy consumption profiles, specific energy consumption, independent parameters [2]. D'Amico A. et al. proposed an artificial neural network to solve the energy and environmental balance buildings [3]. Eom Yong Hwan et al. proposed a novel refrigerant charge fault detection strategy [4]. Jang Juwon et al. analyzed the energy consumption status and over-consumption process to reduce energy consumption according to the results of the analysis [5]. Longo S. et al. evaluated energy efficiency at wastewater treatment plants [6]. Currently some intelligent algorithms are applied to energy saving, such as artificial neural network, convolutional neural network, the artificial neural network and convolutional neural network need a large number of samples to guarantee the sufficient generalization ability, require long learning complexity, and lead to the "over learning" phenomenon. The above problems limit the application of the artificial neural network to some degree, the support vector machine has unique advantage to solve small sample, non-linear and high dimensional pattern recognition problems. The support vector machine is an effective method of energy-saving diagnosis of petrochemical enterprises. The kernel function of support vector machine should satisfy the Mercer theorem, the Gauss kernel function is applied widely to engineering practice.

Yet the kernel function can not closely approximate the curve in $L^2(R^n)$ space, because it is not total orthogonal basis, therefore the support vector machine can not approximate the arbitrary function in classification of it. The wavelet function can establish the orthogonal basis through dilation with translation, which has good approximating ability, therefore the wavelet function can be used to construct wavelet kernel function to improve the generalization ability of support vector machine. In recent years, the wavelet support vector machine has already got application in practice.

Chinnam Siva Koteswara Rao et al. proposed an efficient method to identify unwanted growth in brain using SVM-PUK [7]. Aider Malika Ait et al. was devoted to the off-line handwritten character recognition based on the two dimensional wavelet transform and a single support vector machine classifier [8]. Thirumala Karthik et al. proposed a novel automated recognition based on tunable-Q wavelet transform and a dual multiclass support vector machines [9]. Eseye Abinet Tesfaye et al. proposed a hybrid forecasting model for short-term generation power forecasting [10]. Zhou Ting et al. proposed a data-driven prediction model combining discrete wavelet transform preprocess and support vector machine for groundwater depth forecasting [11]. Garcia Nieto P. J. et al. proposed a new hybrid algorithm, support vector machine with Mexican hat wavelet kernel function for predicting the cyanotoxin content [12]. The advantage of wavelet transform is mainly embodied in analyzing and processing piecewise smooth signal or bounded variance function. The separable wavelet constructed by one dimensional wavelet has only finite orientation when the signal is two or high dimensional function, therefore it can not make optimal use of linear or surface singularity functions with higher dimensions. In order to overcome the

defects of the wavelet transform, Candes and Donoho put forward Curvelet transform and constructed the Curvelet tight frame, the Curvelet transform can realize the stable, efficient and optimal representation of objective function concluding a large number of smooth singular curves. The basic element of Curvelet transform is edge, the basis of it is anisotropic and has strong direction performance. Therefore the Curvelet transform can be combined with support vector machine to construct the Curvelet support vector machine. C. Jimenez-Castaño et al. proposed twin support vector machine for imbalanced data classification [13]. Javad Salimi Sartakhti et al. proposed a fuzzy least squares twin support vector machine that enhances the classification accuracy [14]. Rozhin Eskandarpour and Amin Khodaei proposed twin support vector machine for probabilistic load curtailment estimation [15].

In order to reduce computational complexity and training time, the Curvelet support vector machine can be improved according to basic idea of twin support vector machine. The twin support vector machine is constructed based on idea of approximate support vector machine based on generalized eigenvalue. The history of twin support vector machine is a relatively short time, it has become a hot research topic due to its sound theoretical basis and excellent performance.

The twin Curvelet support vector machine is constructed through combining twin support vector machine and Curvelet support vector machine. The twin Curvelet support vector machine can generate two nonparallel classification hyper planes, making one of the hyper plane as close to a class of samples as possible, and as away from another class of sample. The twin Curvelet support vector machine can transfer single large quadratic programming problem to two small quadratic programming problems, therefore the running time is significantly shortened.

In addition, the parameter selection is a key factor of affecting precision, stability and of twin Curvelet support vector machine. At present, there are three main methods for parameter selection. The first method is experience method, which mainly depends on users and samples, therefore this method has strong subjectivity. The second method is grid search method, its shortcomings lie in the selection of step size, when the step size is small, the amount of calculation is large and the time is long. When step length is long, it is easy to miss the global optimal solution. The third method is intelligent optimization algorithm, which is better than the above two methods. The main intelligent algorithms used for parameter selection of twin Curvelet support vector machine conclude particle swarm algorithm [16], simulated annealing algorithm [17], genetic algorithm [18]. The particle swarm algorithm has the problems of premature convergence and poor local optimization ability. The simulated annealing algorithm falls into local optimum. The programming process of genetic algorithm is complex, the selection of parameters of genetic algorithm is mostly based on experience, and the genetic algorithm is somewhat dependent on selection of initial population. In order to overcome these disadvantages of the above algorithms, a better algorithm should be chosen.

The glowworm swarm optimization is a meta heuristic swarm intelligence optimization algorithm, which has the advantages of clear process, strong robustness and less parameters [19,20]. The glowworm swarm algorithm has some defects, the location updating of it depends on the excellent individuals, if the distance between individuals is too large, glowworm swarm optimization moving speed of firefly with is reduced [21,22]. In order to improve optimization capability of algorithm, the simulated annealing mechanism is used to construct the hybrid glowworm swarm algorithm, and then the defect of falling into local optimum can be avoided. The hybrid glowworm swarm algorithm can improve the global search ability and operation speed,

therefore the hybrid glowworm algorithm is used to optimize the parameters of twin Curvelet support vector machine, then the energy saving level of petrochemical plant can be correctly evaluated based on optimized twin Curvelet support vector machine.

2 Energy saving principle of petrochemical plant

The production process of petrochemical plant mainly concludes transportation, heat exchange, chemical reaction and separation of material involving regulation, control and security of system. All production segments of petrochemical plant have energy usage and consumption, and main energy used concludes fuel gas, electricity, steam and water, and the main process parameters involved conclude pressure, temperature and flow [23]. Efficient and accurate energy saving diagnosis can provide favorable basis for energy saving practice of petrochemical plant. The energy utilization process of petrochemical plant is highly complex, which shows nonlinear, time-varying and random properties. The energy saving diagnosis builds reasonable mathematical model based on wavelet support vector machine according to history data and current data, and then evaluates the energy saving effect of petrochemical plant based on proposed model. The energy saving of petrochemical plant is affected by various factors, the Curvelet support vector machine can be used to carry out energy saving diagnosis of petrochemical plant, which can improve the diagnosis precision. The energy saving diagnosis process of petrochemical plant based on curvelet support vector machine is shown in figure 1.

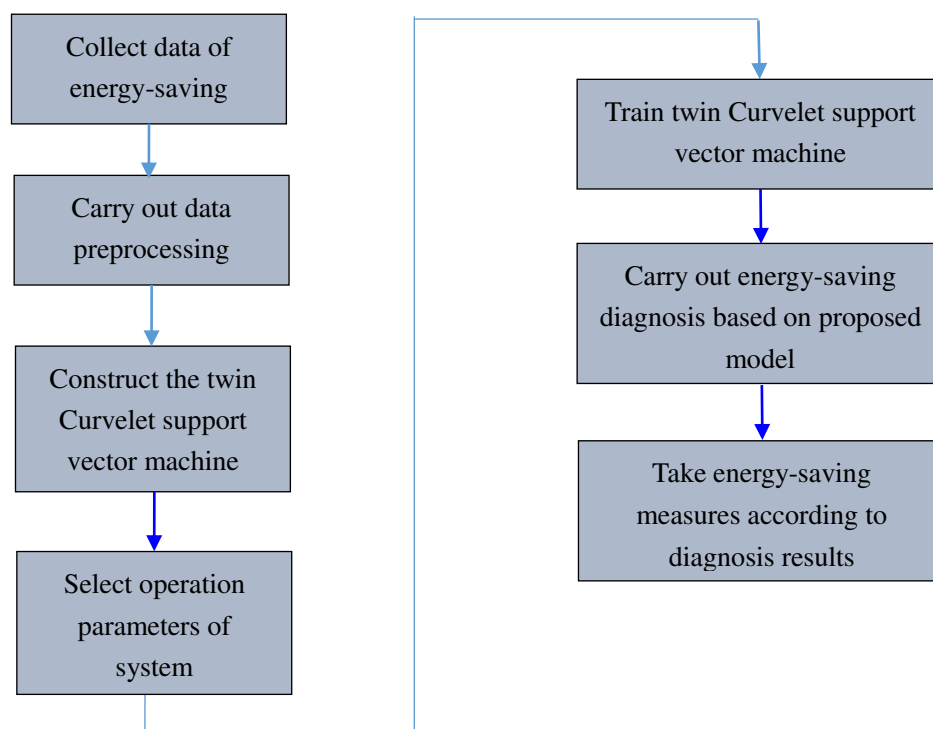


Figure 1 Energy saving diagnosis process of petrochemical plant based on Curvelet support vector machine

3 Construction of twin Curvelet support vector machine

3.1 Mathematical Model

(1) Curvelet kernel function

In the case of two dimensions, the variable in space domain is defined by x , the variable in

frequency domain is defined by w , the polar coordinates in frequency domain are r and θ . Suppose $\varphi_f(x)$ is mother function of Curvelet, and then every Curvelet under j scale can be obtained through rotation and translation of mother function $\varphi_f(x)$. The division amount of angle is defined by A , and the rotation angle is defined by [24]

$$\theta_{j,m} = 2\pi / A \cdot 2^{-[j/2]} \cdot m, \quad m=1,2,\dots, A \cdot 2^{-[j/2]}, \quad 0 < \theta_{j,m} \leq 2\pi \quad (1)$$

where $[j/2]$ is the integral part of $j/2$.

The translation parameter is defined by $k = (k_1, k_2) \in \mathbb{Z}^2$, the Curvelet transform at scale j , direction $\theta_{j,m}$ and position $x_k^{j,m} = R_{\theta_{j,m}}^{-1}(k_1 \cdot 2^{-j}, k_2 \cdot 2^{-j/2})$ is defined by

$$\varphi_{j,m,k}(x) = \varphi_j(R_{\theta_{j,m}}(x - x_k^{j,m})) \quad (2)$$

where $R_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, R_θ is the rotation of θ .

For $f(x) \in L^2(\mathbb{R}^2)$, the Curvelet transform of it can be represented as inner product between it and Curvelet function $\varphi_{j,m,k}(x)$, which is expressed by

$$C(j, m, k) = \langle f, \varphi_{j,m,k} \rangle = \int_{\mathbb{R}^2} f(x) \overline{\varphi_{j,m,k}(x)} dx \quad (3)$$

Given a Curvelet, the tight frame in space $L^2(\mathbb{R}^2)$ can be formed through translation and rotation, and the corresponding reconstruction formula is expressed by [15]

$$f(x) = \sum_{(j,m,k) \in M} \langle f, \varphi_{j,m,k} \rangle \varphi_{j,m,k} \quad (4)$$

where $\varphi_{j,m,k}$ denotes the Curvelet family, M denotes the index collection.

Given the sample set $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$, $x_i \in \mathbb{R}^n$, $y_i \in \mathbb{R}$, l denotes the number of samples, the kernel function of support vector machine is the dot product form $K(x, x') = K(x \cdot x')$, it also is translation invariant form $K(x, x') = K(x - x')$, if a function satisfies Mercer theorem, this function is a support vector machine kernel function.

A necessary and sufficient condition that the symmetric function $K(x, x')$ is the support vector machine kernel function is listed as follows: for all $g(x) \neq 0$ that satisfies

$$\int_{\mathbb{R}^n} g^2(x) dx < \infty, \text{ the following inequality holds:}$$

$$\iint K(x, x')g(x)g(x')dxdx' \geq 0, x, x' \in R^n \quad (5)$$

If the Fourier transform satisfies the following condition:

$$F(\omega) = (2\pi)^{-n/2} \int_{R^n} e^{-j\omega x} K(x)dx \geq 0, x \in R^n \quad (6)$$

the translation invariant form $K(x, x') = K(x - x')$ satisfying condition (5) is a permissible support vector machine kernel function.

The Curvelet kernel function is expressed by

$$K(x, x') = \prod_{i=1}^l \varphi_{j,m,k}(\frac{x_i - x'_i}{a}) \quad (7)$$

where a is the Curvelet scale parameter, $a > 0$.

(2) Mathematical model of twin Curvelet support vector machine

The Curvelet function is combined with twin support vector machine to construct the twin Curvelet support vector machine. The twin Curvelet support vector machine divide the training set into two parts, and then two non parallel decision functions can be constructed through selecting proper loss function for the two parts respectively.

The following training set is given: $S = \{ (x_1, y_1), (x_2, y_2), \dots, (x_m, y_m) \}$, where x_1, x_2, \dots, x_m denotes the data input, y_1, y_2, \dots, y_m denotes the data output. The matrix $A \in R^{m \times n}$ and output vector $\mathbf{y} = [y_1, y_2, \dots, y_m]^T$ are used to denote m inputting samples in n dimensional space R , the i th line of A is defined by x_i^T , where m is number of samples of training set.

The two decision functions of twin Curvelet support vector machine are not parallel, which are expressed by

$$f_1(\mathbf{x}) = (\mathbf{w}_1^T, \mathbf{x}) + b_1 \quad (8)$$

$$f_2(\mathbf{x}) = (\mathbf{w}_2^T, \mathbf{x}) + b_2 \quad (9)$$

where $\mathbf{x} \in R^n$, \mathbf{w}_1 and \mathbf{w}_2 are the weight vectors, b_1 and b_2 are offsets, (\cdot, \cdot) denotes the inner product.

The original problems are expressed by

$$\min_{\mathbf{w}_1, b_1, \zeta} \left(\frac{1}{2} \left\| \mathbf{y} - (\mathbf{A}\mathbf{w}_1 + \mathbf{e}b_1) - \mathbf{e}\varepsilon_1 \right\|^2 + C_1 \mathbf{e}^T \zeta \right) \quad (8)$$

$$\text{s.t. } \mathbf{y} - (\mathbf{A}\mathbf{w}_1 + \mathbf{e}b_1) \geq \mathbf{e}\varepsilon_1 - \zeta, \quad \zeta \geq 0$$

$$\min_{\mathbf{w}_2, b_2, \zeta} \left(\frac{1}{2} \left\| \mathbf{y} - (\mathbf{A}\mathbf{w}_2 + \mathbf{e}b_2) - \mathbf{e}\varepsilon_2 \right\|^2 + C_2 \mathbf{e}^T \zeta \right) \quad (9)$$

$$\text{s.t. } (\mathbf{A}\mathbf{w}_2 + \mathbf{e}b_2) - \mathbf{y} \geq \mathbf{e}\varepsilon_2 - \zeta, \quad \zeta \geq 0$$

where $\|\cdot\|$ is L_2 norm, C_1 and C_2 are penalty coefficients, ε_1 and $\varepsilon_2 > 0$ are constants,

ξ and β are relaxation vectors, e is $m \times 1$ dimensional unit column vector.

The dual forms of expressions (8) and (9) can be obtained based on Lagrangian function, which is expressed by [16]

$$\min_{\alpha} \frac{1}{2} \alpha^T S(S^T S)^{-1} S^T \alpha - d_1^T S(S^T S)^{-1} S^T \alpha + d_1^T \alpha \quad (10)$$

$$\text{s.t. } 0 \leq \alpha \leq C_1 e$$

$$\min_{\alpha} \frac{1}{2} \beta^T S(S^T S)^{-1} S^T \alpha - d_2^T S(S^T S)^{-1} S^T \beta + d_2^T \beta \quad (11)$$

$$\text{s.t. } 0 \leq \beta \leq C_2 e$$

where $S = \begin{bmatrix} A \\ e \end{bmatrix}$, $d_1 = y - \varepsilon_1 e$, $d_2 = y + \varepsilon_2 e$.

The optimal solutions of equations (10) and (11) are listed as follows:

$$[\alpha_1^T \ b_1]^T = (S^T S)^{-1} S^T (d_1 - \alpha) \quad (12)$$

$$[\alpha_2^T \ b_2]^T = (S^T S)^{-1} S^T (d_2 + \beta) \quad (13)$$

In the case of non linearity, the solutions can be obtained through mapping Curvelet kernel function to high dimensional space, and the twin Curvelet support vector machine decision functions are expressed by

$$f_1(x) = K(x^T, A^T) \alpha_1 + b_1 \quad (14)$$

$$f_2(x) = K(x^T, A^T) \alpha_2 + b_2 \quad (15)$$

When $f_1(x)$ and $f_2(x)$ are solved, the decision function is expressed by

$$f(x) = \frac{f_1(x) + f_2(x)}{2} \quad (16)$$

(3) Training algorithm of twin Curvelet support vector machine

The twin Curvelet support vector machine has three parameters concluding C_1 , C_2 and a , the three parameters are closely related with diagnosis precision, therefore an effective optimal algorithm is required to obtain the best parameters of twin Curvelet support vector machine.

(a) Mathematical model of improved glowworm swarm optimization

The glowworm swarm optimization (GSO) algorithm is a novel intelligent optimization algorithm, which already has been applied in many fields, but there has been very few successful application of GSO in parameter optimization of support vector machine. The traditional GSO algorithm is easy to fall into local optimum and convergence insufficiency, and it has low population diversity. In order to improve the performance of GSO and local search ability, therefore the GSO algorithm is improved based on dynamic step [18].

The GSO algorithm is used to simulate the mutual attraction and location update between glowworms. The location and the fluorescein in the location of i th glowworm are defined by $x_i(t)$ and $y_i(t)$ respectively, and every location corresponds to a objective function $f(x_i(t))$.

The location of i th glowworm is occurs randomly according to following expression:

$$x_i = L + rad(0,1) \cdot (U - L), \quad i = 1, 2, \Lambda, M \quad (15)$$

where M denotes the population size of glowworms, L denotes the lower bound of x_{ij} , U denotes the upper bound of x_{ij} , rad denotes the random number yielding uniform distribution on domain (0,1).

When every glowworm transmit information to neighboring individuals in decision range of its region, the decision range is updated by [19]

$$r_d^i(t+1) = \min\{r_s, \max[0, r_d^i(t) + \beta(n_t - |N_i(t)|)]\} \quad (16)$$

where $r_d^i(t)$ denotes the decision radius in region of i th glowworm in t th iteration, $0 < r_d^i(t) < r_s$; r_s denotes the maximum perception radius, n_t denotes the threshold of individual quantity in controlling neighborhood, β denotes the variation ratio of neighborhood, $N_i(t)$ denotes the group of individuals in neighborhood of i th glowworm in the t th iteration.

The individual quantity of i th glowworm in decision range is calculated by

$$N_i(t) = \{j : \|x_j(t) - x_i(t)\| < r_d^i(t); l_i(t) < l_j(t)\} \quad (17)$$

where $x_j(t)$ denotes the location of j th glowworm, $l_i(t)$ denotes the fluorescein value.

The motion direction of i th glowworm is decided by number of fluorescein of every glowworm in the neighborhood. The probability of i th glowworm moving to j th glowworm is calculated by [20]

$$p_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)} \quad (18)$$

The new location of i th glowworm after movement is calculated by

$$x_i(t+1) = x_i(t) + \alpha e^{-p_{ij} r_{ij}} (x_j(t) - x_i(t)) \quad (19)$$

where r_{ij} denotes the distance between i th glowworm and j th glowworm, α denotes the step length factor.

After the location of i th glowworm is updated, the new fluorescein value is calculated by

$$l_i(t) = (1 - \lambda) l_i(t-1) x_i(t) + \eta f(x_i(t)) \quad (20)$$

where λ denotes the constant, $0 < \lambda < 1$, which is related to volatilisation of fluorescein; η denotes the constant, which can identify the update rate of fluorescein.

In order to improve performance of GSO algorithm, the adaptive search step is introduced into the traditional GSO algorithm. The step length factor α is regulated by

$$\alpha(t+1) = 0.85\alpha(t) \quad (21)$$

$$\alpha(0) = 0.5(x_{d,\max} - x_{d,\min}) \quad (22)$$

where $x_{d,\max}$ and $x_{d,\min}$ are upper bound and lower bound of objective function.

In order to improve the search velocity of algorithm, the regulating formulas of λ and η are respectively expressed by [21]

$$\lambda(t+1) = (1 - \frac{t}{N_{\max}^2})\lambda(t) \quad (23)$$

$$\eta(t+1) = (1 - \frac{t}{N_{\max}^2})\eta(t) \quad (24)$$

The GSO algorithm will fall into local optimum when the individual searches the optimal value for a long time within the radius of neighborhood, and then fitness degree function does not change. Therefore the chaos seeking strategy is used to generate new solution to optimize the individuals, and then the global optimization ability of algorithm is improved. The Chebyshev chaotic mapping is used as search strategy, which is expressed by [22]

$$y_{k+1} = \cos(k \arccos y_k) \quad (25)$$

The location is updated by

$$X_i^{m'} = X_i^m + y_k \quad (26)$$

where m denotes the spatial dimension, $k=1,2,\Lambda, Max$, X_i^m denotes the original location,

$X_i^{m'}$ denotes the optimized location. The fitness degree function value of new individual is

calculated based on $X_i^{m'}$, which is compared with current stagnant individual, if the fitness degree of the new individual is better than that of stagnant individual, the stagnant individual is replaced by the new individual.

(b) Hybrid glowworm swarm algorithm based simulated annealing algorithm

The simulated annealing algorithm is established based on the principle of solid annealing, the solid is heated to enough high temperature and then is cooled slowly. According to Metropolis criterion, the probability that particles tend to equilibrium at temperature T is $e^{-\Delta E/(kT)}$, where E is the internal energy at temperature T , ΔE is the changing amount, k is Boltzmann

constant. Using solid annealing to simulated combined optimization problem, the internal energy E is simulated as the objective function value, the temperature T is evolved into controlling parameter, and then the simulated annealing algorithm can be obtained. The simulated annealing algorithm can be used to construct the hybrid glowworm swarm algorithm.

Firstly, the selection mechanism is improved. In order to dynamically adjust the selection pressure in the process of searching the optimal solution, the Boltzmann selection strategy is introduced into the search process of glowworm swarm algorithm. $P_{ij}(t)$ denotes the probability that the i th glowworm moves to the j th glowworm in its neighborhood set at time t .

$$P_{ij} = \frac{e^{\frac{l_j(t)-l_i(t)}{T}}}{\sum_{k \in N_i(t)} e^{\frac{l_j(t)-l_i(t)}{T}}} \quad (27)$$

where $T = T_0 \alpha^{C-1}$, T denotes the temperature in simulated annealing algorithm, T_0 denotes the initial temperature, α denotes the decay factor, C denotes the number of iterations.

In improved glowworm swarm algorithm the obtained optimal solution is often approximate value due to the fixed moving step size and randomness of selecting a neighbor of artificial glowworm, therefore it is difficult to obtain the accurate solution further. Therefore the strong local search ability of simulated annealing algorithm can be used to improve the accuracy of solving the problem for improved glowworm swarm algorithm. The flow of hybrid glowworm swarm algorithm is listed as follows:

Step 1: The parameters of improved glowworm swarm algorithm concluding $r_d^i(t)$, r_s , n_t , β , the glowworm population size N , the number of iterations $t=0$, the maximum iteration times L_{\max} , and the parameters of simulated annealing algorithm concluding T_0, α , iteration times L at each temperature are initialized.

Step 2: The initial position of each glowworm is randomly initialized, and the objective function value of each glowworm that is fitness degree is calculated, and then the bulletin board is initialized to record the best glowworm individual.

Step 3: The fluorescein values of all glowworms are calculated based on equation (20).

Step 4: In the moving stage, the neighbor glowworm set $N_i(t)$ in the decision domain of each glowworm is calculated according to equation (17).

Step 5: For any glowworm i , the probability $P_{ij}(t)$ of its moving to each neighbor is calculated according to equation (27). A neighbor is selected according to the probability to move, and then location of it is updated based on equation (26).

Step 6: The decision domain radius of each glowworm is updated based on equation (16).

Step 7: The glowworm $best$ with best fitness degrees in this generation is found out, and the corresponding location x_{best} is recorded. If it is better than bulletin board information, and the

bulletin board information will be updated.

Step 8: $t \leftarrow t + 1$, an iteration is completed, judge whether the condition $t < L_{\max}$ is satisfied.

If the condition is met, return to step 3 and enter the next iteration. Otherwise exit the iteration and output the result.

4 Case study

In order to verify the effectiveness of proposed energy saving diagnosis, a LPG mercaptan removal plant in a oil refinery is used as research object to carry out numerical analysis. Merox mercaptan removal technology is applied in this plant, and the LPG yield reaches about 19% after transformation. The plant has a large energy consumption, it is necessary to carry out energy-saving diagnosis for taking effective energy saving measures. The diagnostic index system of energy saving is confirmed according to practical situation of plant, which is listed in table 1.

Table 1 Diagnostic index system of energy saving of LPG mercaptan removal plant

Type	Index
Industrial residue	Sulfur content of raw oil/%
	Composition of LPG (φ)/%
	Yield of LPG/ $t \cdot h^{-1}$
	Mass concentration of sulfur before mercaptan removal/ $mg \cdot m^{-3}$
	Mass concentration of sulfur after mercaptan removal/ $mg \cdot m^{-3}$
	Consumption of fresh lye/ $t \cdot a^{-1}$
	Emission of alkaline residue/ $t \cdot a^{-1}$
Waste water	Mass fraction of VOCs/ $\mu g \cdot g^{-1}$
	PH of the wash water
	Consumption of desalinated water/ $t \cdot h^{-1}$
Waste gas	Component of tail gas/%
	Operating pressure of regeneration unit/MPa

The Curvelet support vector machine optimized by improved GSO algorithm is applied to diagnose the energy-saving effect of the LPG mercaptan removal plant after the index system is confirmed. The energy-saving effect of the LPG mercaptan removal plant is measured by energy consumption.

The initial parameters of improved GSO algorithm are listed as follows: $N=80$, $r_d^i(t)=9$,

$r_s=15$, $n_t=6$, $\beta=0.07$, $\alpha \in [0.01,1]$, $\lambda=0.55$, $\eta=0.45$, $L_{\max}=250$. The parameters of

simulated annealing algorithm are listed as follows: $T_0=120$, $\alpha=0.90$, $L=60$.

The data collected for 24 months from 2018 to 2019 is collected from the oil refinery, which is used as training sample. The data collected from January to October in 2020 is used as testing sample. The support vector machine optimized by simulated annealing algorithm (SVM-SA), the twin B-spline wavelet support vector machine optimized by traditional GSO algorithm (TBWSVM-TGSO), and the twin Curvelet support vector machine trained by improved hybrid glowworm swarm algorithm (TCSVM-HGSO) are applied to prediction simulation analysis. The mean absolute percentage error (MAPE) and determination coefficient R^2 are used to evaluate the prediction effect of different models.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|\hat{E} - E|}{E} \times 100\% \quad (27)$$

$$R^2 = \frac{\left[N \sum_{i=1}^N \hat{E} \cdot E - \sum_{i=1}^N \hat{E} \cdot \sum_{i=1}^N E \right]^2}{\left[N \sum_{i=1}^N \hat{E}^2 - \left(N \sum_{i=1}^N \hat{E} \right)^2 \right] \cdot \left[N \sum_{i=1}^N E^2 - \left(\sum_{i=1}^N E \right)^2 \right]} \quad (28)$$

where \hat{E} denotes the diagnostic value, E denotes the actual value, N denotes the number of samples.

The comparison results of predicted and actual values of training samples for different models are shown in figure 1 to figure 3. Figure 1 shows the comparison results of SVM-SA, and the determination coefficient R^2 equals 0.925. Figure 2 shows the comparison results of TBWSVM-TGSO, the determination coefficient R^2 equals 0.936, the prediction precision of TBWSVM-TGSO is better than that of SVM-GA. Figure 3 shows the comparison results of TCSVM-HGSO, the determination coefficient R^2 equals 0.945, results show that the TCSVM-HGSO model has best fitting precision among the three models, and can obtain the best prediction precision.

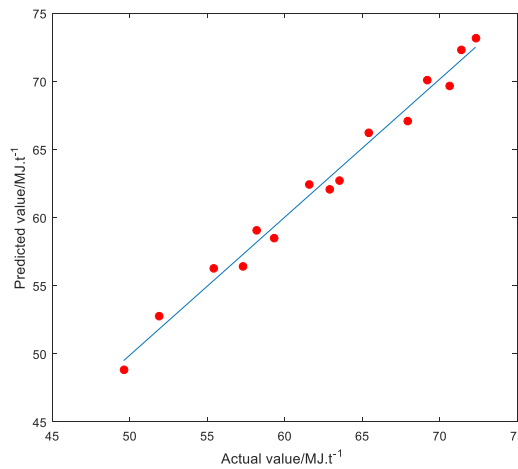


Figure 1 Predicted results of SVM-SA ($R^2=0.925$)

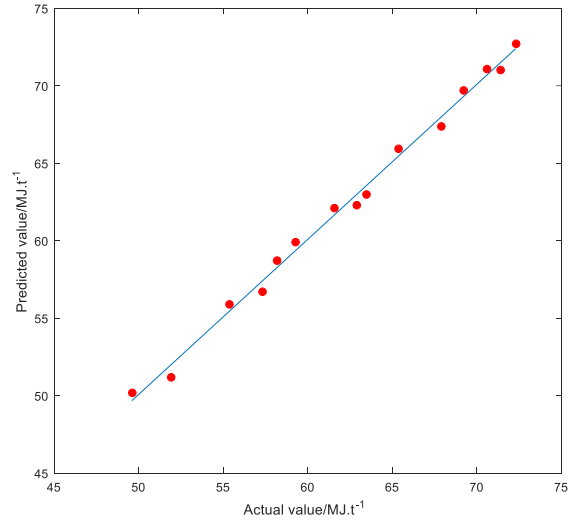


Figure 2 Predicted results of TBWSVM-TGSO ($R^2=0.936$)

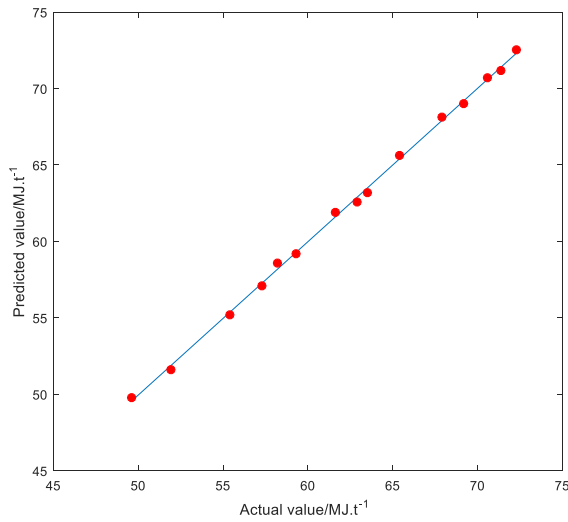


Figure 3 Predicted results of TCSVM-HGSO ($R^2=0.945$)

The optimal parameters C_1 , C_2 and a are obtained based on three models, which are listed in table 2. The optimal parameters are used to carry out prediction simulation for testing samples.

Table 2 Optimal parameters C_1 , C_2 and a based on different models

Prediction model	C_1	C_2	a
SVM-GA	7.82	6.94	0.12
TBWSVM-TGSO	5.62	4.95	0.22
TCSVM-HGSO	4.57	3.96	0.39

The predict simulation of testing samples are carried out based on the three models, and prediction results are shown in figure 4, results show that the changing trend of energy consumption of LPG mercaptan removal plant. The predicted value of energy consumption by the TCSVM-HGSO is closer to actual value than other models, results show that the proposed TCSVM-HGSO has better prediction accuracy than other models.

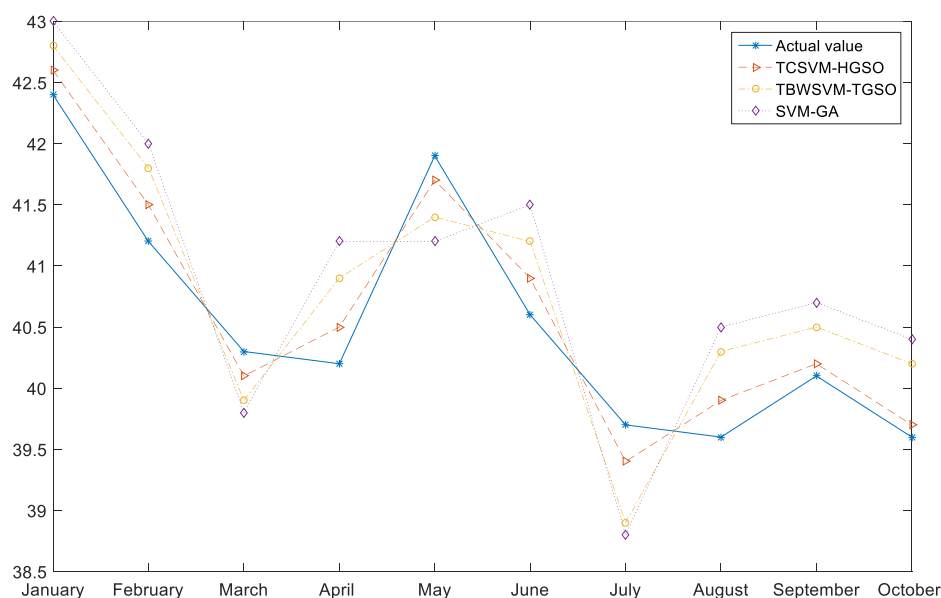


Figure 4 Prediction results of testing samples based on different models

The mean absolute percentage error (MAPE) of different predicted models are shown in figure 5. As seen from figure 5, the MAPE of TCSVM-HGSO is less than that of other two models, therefore the TCSVM-HGSO has best prediction precision. The main reason for this change lies in that the TCSVM-HGSO can better break away from the local convergence and avoid over-fitting.

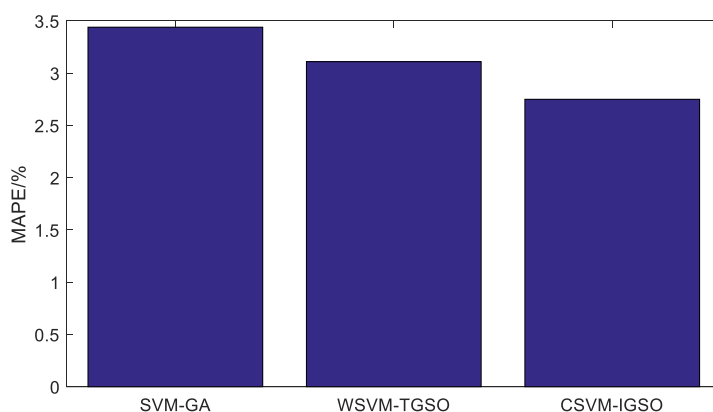


Figure 5 MAPE of different prediction models

5 Conclusions

Energy-saving diagnosis is important for improving economic benefit of petrochemical plant, therefore the curvelet transform is combined with the twin support vector machine to construct the twin curvelet support vector machine, in order to improve the prediction effectiveness of the twin curvelet support vector machine, the hybrid glowworm swarm algorithm (HGSO) is established

based on simulated annealing algorithm to optimize the parameters of twin Curvelet support vector machine. A LPG mercaptan removal plant in a oil refinery is used as the research objective to carry out energy-saving diagnosis simulation based on SVM-GA, TBWSVM-TGSO and TCSVM-HGSO, the simulation results of training samples show that determination coefficient R^2 of the proposed TCSVM-HGSO is highest, the prediction value of TCSVM-HGSO is closest to actual value, therefore TCSVM-HGSO has the highest diagnosis precision. The simulation results of testing samples show that the mean absolute percentage error (MAPE) of TCSVM-HGSO is least, and TCSVM-HGSO can correctly reflect the the changing trend of energy consumption of LPG mercaptan removal plant. Therefore the TCSVM-HGSO has a good respect in energy-saving of petrochemical plant.

Ethical approval

No

Funding details (In case of Funding)

No

Conflict of interest

No

Informed Consent

All authors read and approve the final manuscript.

Author's contribution

B is in charge of the model construction, wrote and revised the manuscript. D takes part in the program design. DK takes part in data analysis. LZ takes part in simulation analysis. All authors read and approve the final manuscript.

References

- [1] Wantao Shi, Xinqiao Jin, Yijun Wang, Evaluation of energy saving potential of HVAC system by operation data with uncertainties, Energy and Buildings, Volume 204, 2019, Article 109513.
- [2] Kasturi Srikant and Kannappan Bakthavatsalam Ayalur, Energy diagnosis and savings potential in knitted fabric based wet processing units, Energy for Sustainable Development, Vol.54, 2020, pp.85-100.
- [3] D'Amico A., Ciulla G., Traverso M., Lo Brano V., Palumbo E., Artificial Neural Networks to assess energy and environmental performance of buildings: An Italian case study, Journal of Cleaner Production, Vol.239, 2019, Article 117993.
- [4] Eom Yong Hwan, Yoo Jin Woo, Hong Sung Bin, Kim Min Soo, Refrigerant charge fault detection method of air source heat pump system using convolutional neural network for energy saving, Energy, Vol.187, 2019, Article 115877.
- [5] Jang Juwon, Park Seokhun, Park Jongseok, Kim Hoseop, Kim Jaemoon, Lee Hosik, Energy assessment and savings for the sewage treatment plant using EUAT and Tool-kit in Korea, Desalination And Water Treatment, Vol.154, 2019, pp.132-140.
- [6] Longo S., Mauricio-Iglesias M., Soares A., Campo P., Fatone F., Eusebi A.L., Akkersdijk E., Stefani L., Hospido A., ENERWATER - A standard method for assessing and improving the energy efficiency of wastewater treatment plants, Applied Energy, Vol.242, 2019, pp.897-910.

- [7] Chinnam Siva Koteswara Rao, Sistla Venkatramaphanikumar, Kolli Venkata Krishna Kishore, SVM-PUK Kernel Based MRI-brain Tumor Identification Using Texture and Gabor Wavelets, *Traitement Du Signal*, Vol.36, No.2, 2019, pp.185-191.
- [8] Aider Malika Ait, Hammouche Kamal, Gaceb Djamel, Recognition of Handwritten Characters Based on Wavelet Transform and SVM Classifier, *International Arab Journal Of Information Technology*, Vol.15, No.6, 2018, pp.1082-1087.
- [9] Thirumala Karthik, Prasad M. Siva, Jain Trapti, Umarikar Amod C., Tunable-Q Wavelet Transform and Dual Multiclass SVM for Online Automatic Detection of Power Quality Disturbances, *IEEE Transactions On Smart Grid*, Vol.9, No.4, pp.3018-3028.
- [10] Eseye Abinet Tesfaye, Zhang Jianhua, Zheng Dehua, Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information, *Renewable Energy*, Vol.118, 2018, pp.357-367.
- [11] Zhou Ting, Wang Faxin, Yang Zhi, Comparative Analysis of ANN and SVM Models Combined with Wavelet Preprocess for Groundwater Depth Prediction, *Water*, Vol.9, No.10, Article 781.
- [12] Garcia Nieto P.J., Garcia-Gonzalo E., Alonso Fernandez J.R., Diaz Muniz C., A hybrid wavelet kernel SVM-based method using artificial bee colony algorithm for predicting, the cyanotoxin content from experimental cyanobacteria concentrations in the Trasona reservoir (Northern Spain), *Journal of Computational and Applied Mathematics*, VOL.309, 2017, pp.587-602.
- [13] C. Jimenez-Castaño, A. Alvarez-Meza, A. Orozco-Gutierrez. Enhanced automatic twin support vector machine for imbalanced data classification, *Pattern Recognition*, 2020, 107(11):107442.
- [14] Javad Salimi Sartakhti, Homayun Afrabandpey, Nasser Ghadiri. Fuzzy least squares twin support vector machines, *Engineering Applications of Artificial Intelligence*, 2019, 85(10):402-409.
- [15] Rozhin Eskandarpour and Amin Khodaei, Probabilistic load curtailment estimation using posterior probability model and twin support vector machine, *Journal Modern Power Systems and Clean Energy*, 2019, 7(4):665-675.
- [16] Jiaohui Xu, Wen Tan, Tingshun Li, Predicting fan blade icing by using particle swarm optimization and support vector machine algorithm, *Computers & Electrical Engineering*, 2020, 87:106751.
- [17] Farzaneh Sajedi Hosseini, Bahram Choubin, Amir Mosavi, Narjes Nabipour, Shahaboddin Shamshirband, Hamid Darabi, Ali Torabi Haghighi, Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: Application of the simulated annealing feature selection method, *Science of The Total Environment*, 2020, 711:135161.
- [18] Rahul, Bharat Choudhary, An Advanced Genetic Algorithm with Improved Support Vector Machine for Multi-Class Classification of Real Power Quality Events, *Electric Power Systems Research*, 2021, 191: 106879.
- [19] Yongjun Sun, Rui Ma, Jiayu Chen, Tao Xu, Heuristic optimization for grid-interactive net-zero energy building design through the glowworm swarm algorithm, *Energy and Buildings*, 2020, Vol.208, February, 109644.
- [20] Jyotika Pruthi, Kavita Khanna, Shaveta Arora, Optic Cup segmentation from retinal fundus images using Glowworm Swarm Optimization for glaucoma detection, *Biomedical Signal*

Processing and Control, 2020, 60:102004.

[21] Ali Khan, Farooq Aftab, Zhongshan Zhang, Self-organization based clustering scheme for FANETs using Glowworm Swarm Optimization, *Physical Communication*, 2019, 36:100769.

[22] Aparajita Chowdhury, Debashis De, MSLG-RGSO: Movement score based limited grid-mobility approach using reverse Glowworm Swarm Optimization algorithm for mobile wireless sensor networks, *Ad Hoc Networks*, 2020, 106:102191.

[23] Alireza Talaei, Md. Ahiduzzaman, Amit Kumar, Assessment of long-term energy efficiency improvement and greenhouse gas emissions mitigation potentials in the chemical sector, *Energy*, Vol.153, 2018, pp.231-247.

[24] Hasan Al-Marzouqi, Yuting Hu, Ghassan AlRegib, Texture retrieval using periodically extended and adaptive curvelets, *Signal Processing: Image Communication*, Vol.76, 2019, pp.252-260.

[25] Deepika Sharma, Kavita Goyal, Rohit Kumar Singla, A curvelet method for numerical solution of partial differential equations, *Applied Numerical Mathematics*, Vol.148, 2020, pp.28-44.

[26] Julio López, Sebastián Maldonado, Miguel Carrasco, Robust nonparallel support vector machines via second-order cone programming, *Neurocomputing*, Vol.364, 2019, pp.227-238.

[27] Lamei Shi, Jiahua Zhang, Da Zhang, Tertsea Igbawua, Yuqin Liu, Developing a dust storm detection method combining Support Vector Machine and satellite data in typical dust regions of Asia, *Advances in Space Research*, Vol.65, No.4, 2020, pp.1263-1278.

[28] Puttamadappa C. and Parameshachari B.D., Demand side management of small scale loads in a smart grid using glow-worm swarm optimization technique, *Microprocessors and Microsystems*, Vol.71, 2019, 102886.

[29] Hasan Koyuncu, Rahime Ceylan, A PSO based approach: Scout particle swarm algorithm for continuous global optimization problems, *Journal of Computational Design and Engineering*, Vol.6, No.2, 2019, pp.129-142.

[30] Utkarsh Goel, Shubham Varshney, Anshul Jain, Saumil Maheshwari, Anupam Shukla, Three Dimensional Path Planning for UAVs in Dynamic Environment using Glow-worm Swarm Optimization, *Procedia Computer Science*, Vol.133, 2018, pp.230-239.

[31] Saeed Balochian and Hossein Baloochian, Social mimic optimization algorithm and engineering applications, *Expert Systems with Applications*, Vol.134, 2019, pp.178-191.

[32] Manjit Kaur, Dilbag Singh, Kehui Sun, Umashankar Rawat, Color image encryption using non-dominated sorting genetic algorithm with local chaotic search based 5D chaotic map, *Future Generation Computer Systems*, Vol.107, 2020, pp.333-350.

Figures

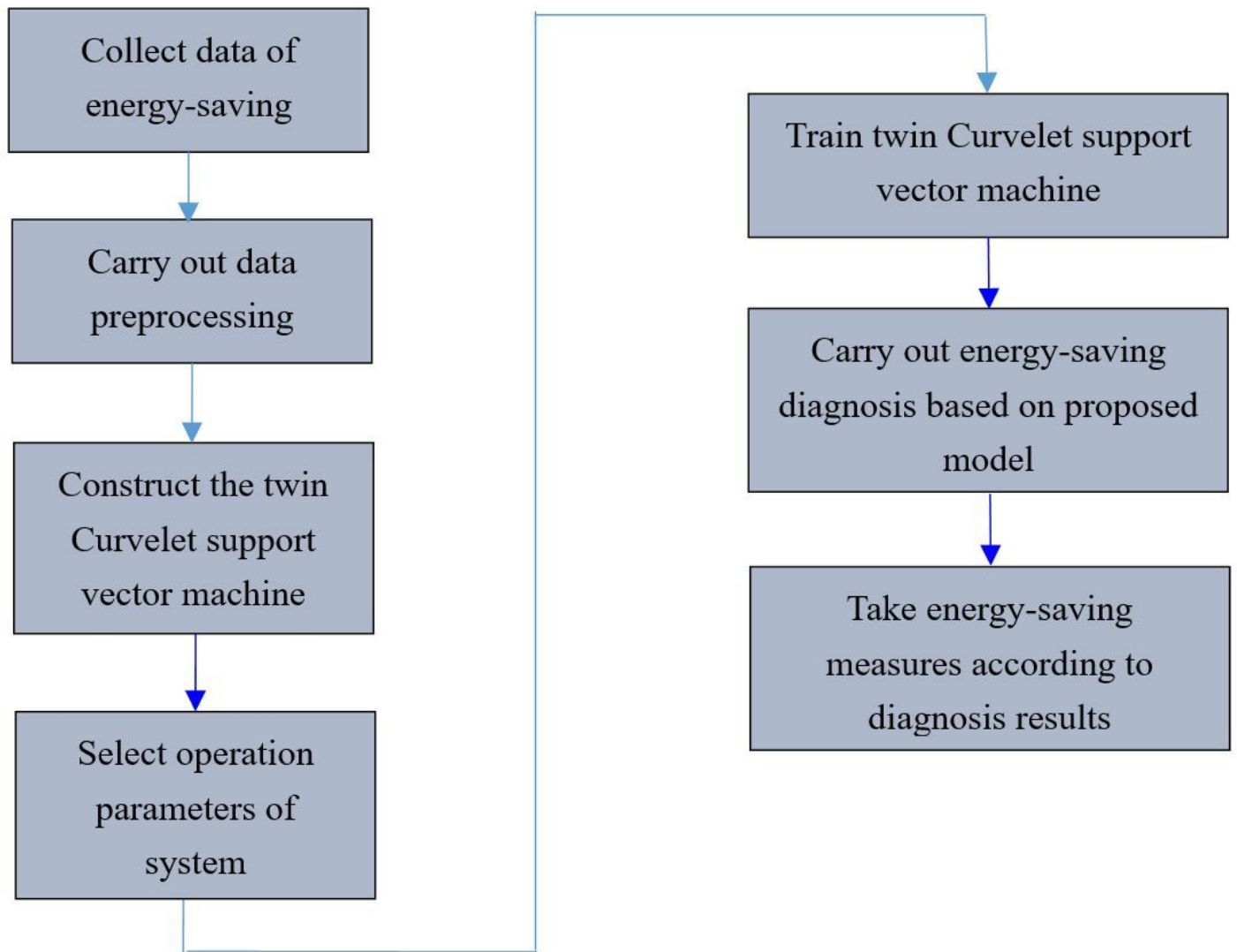


Figure 1

Energy saving diagnosis process of petrochemical plant based on Curvelet support vector machine

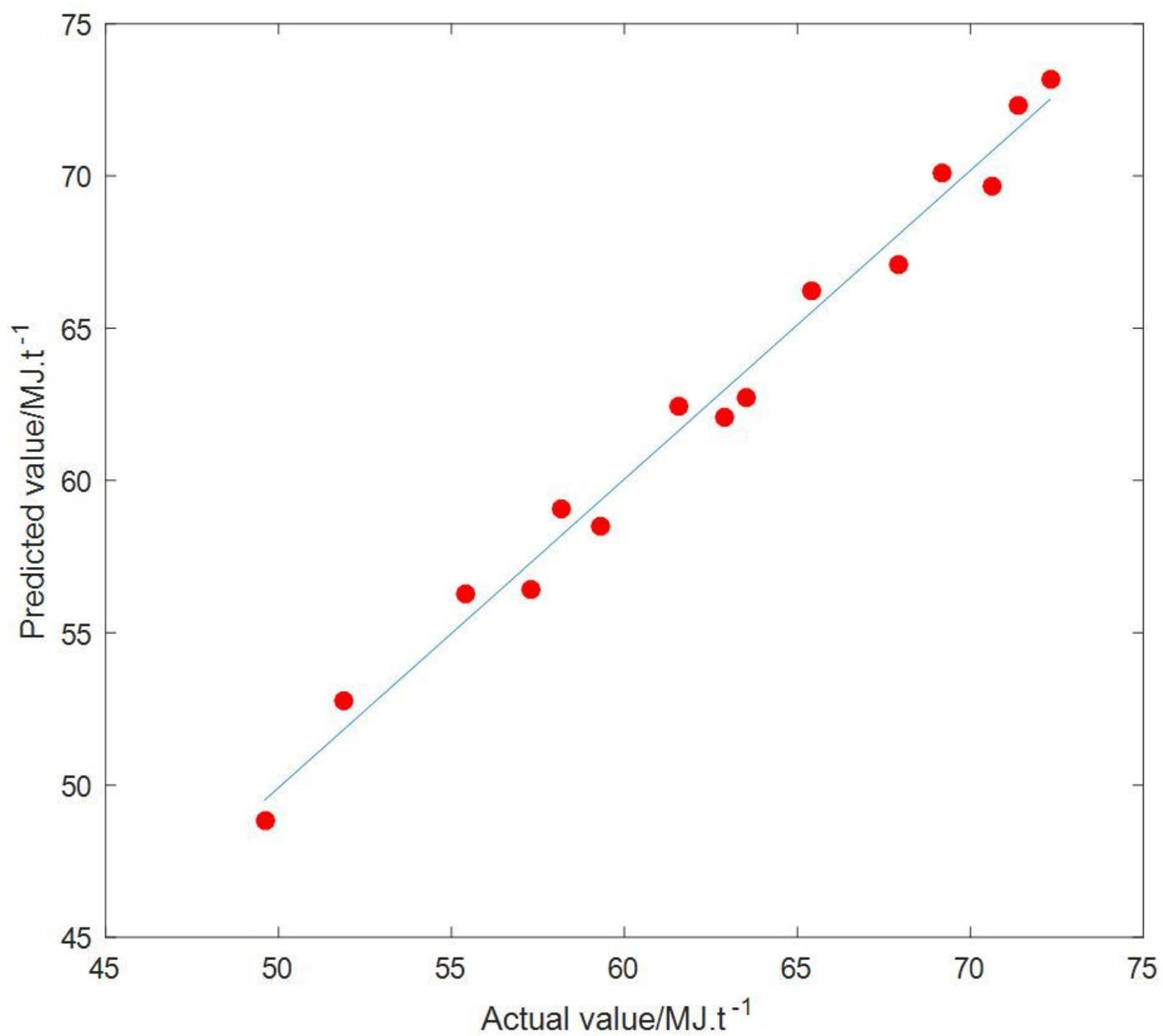


Figure 2

Predicted results of SVM-SA (R2=0.925)

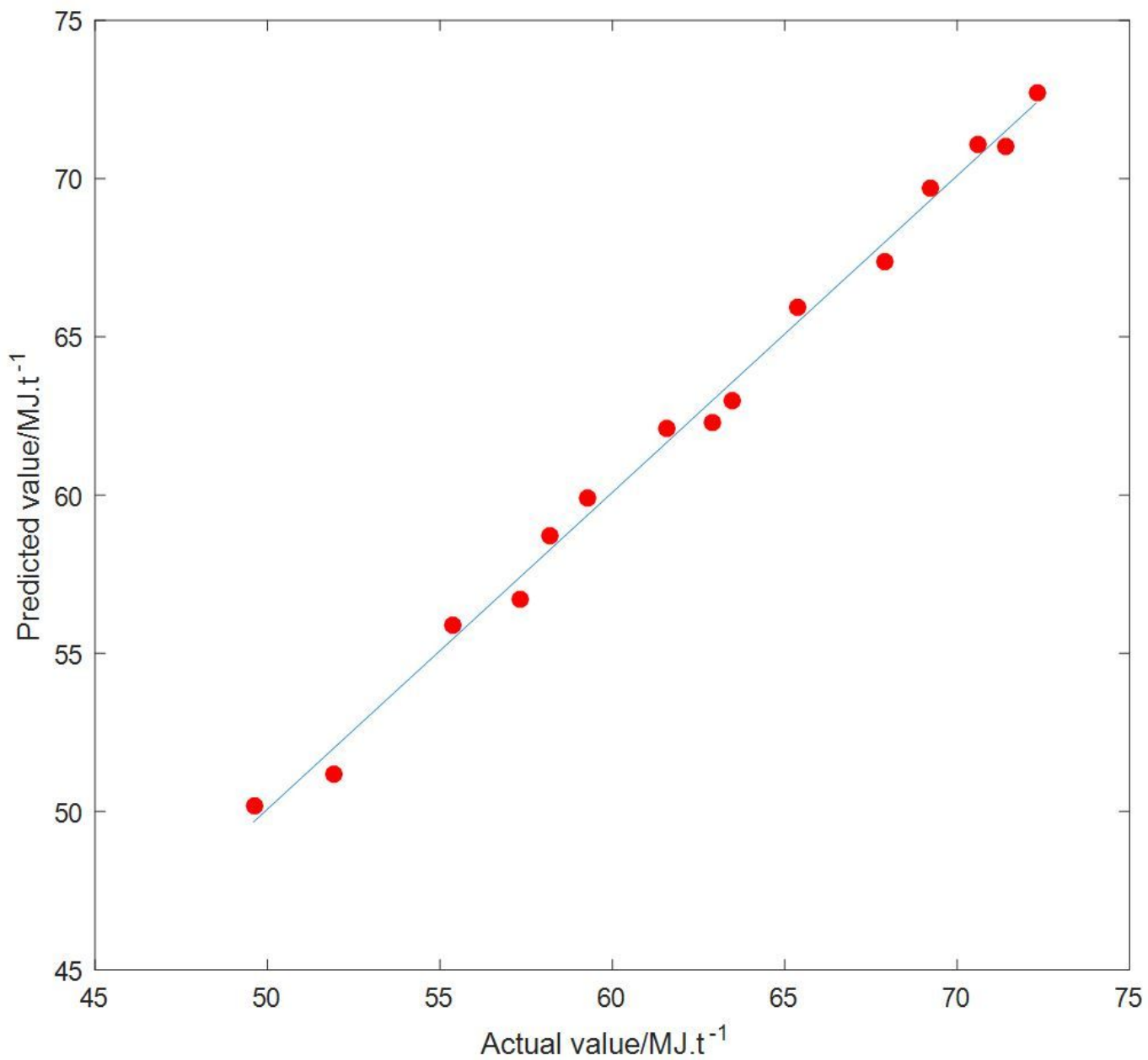


Figure 3

Predicted results of TBWSVM-TGSO (R3=0.936)

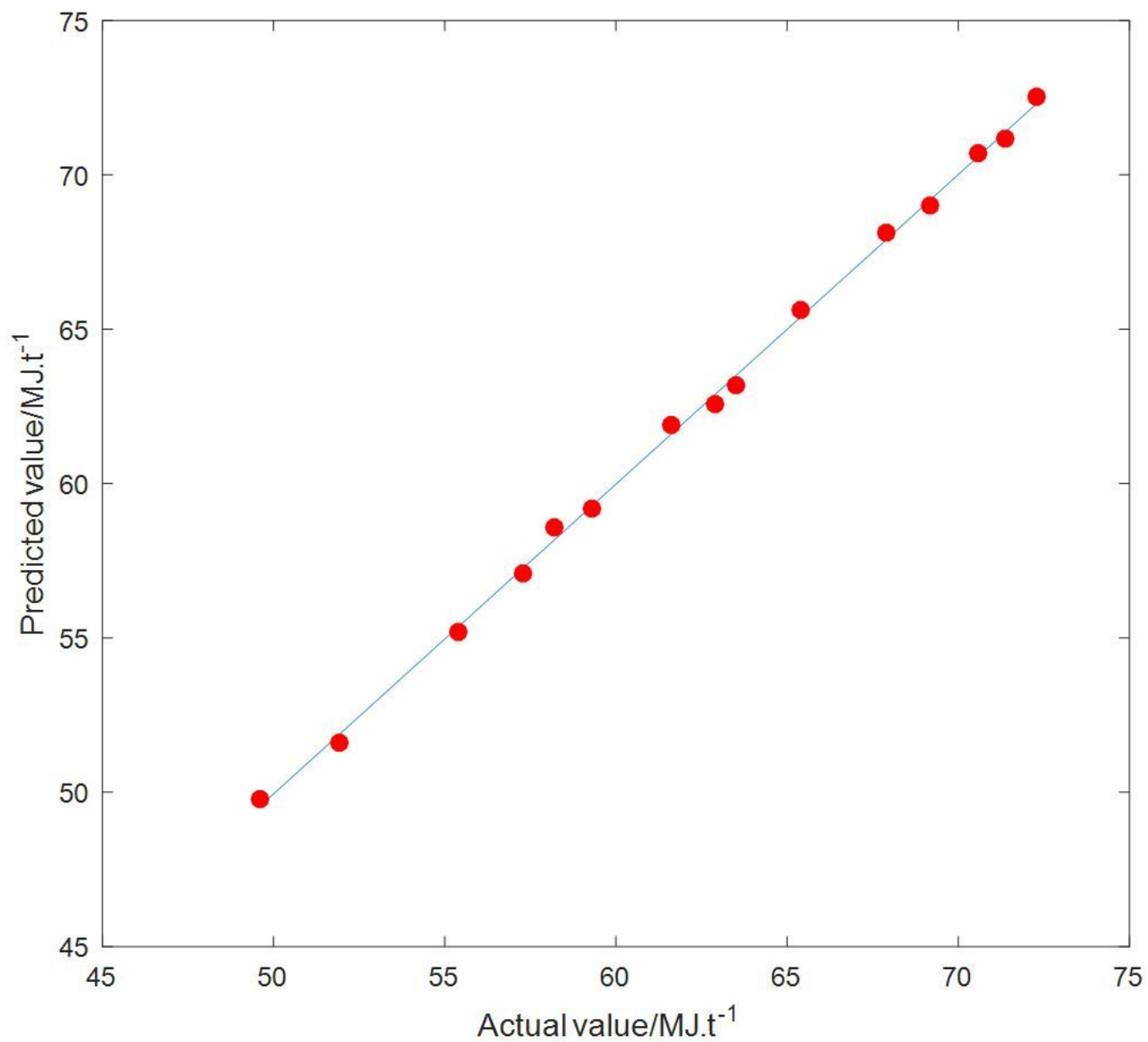


Figure 4

Predicted results of TCSVM-HGSO (R²=0.945)

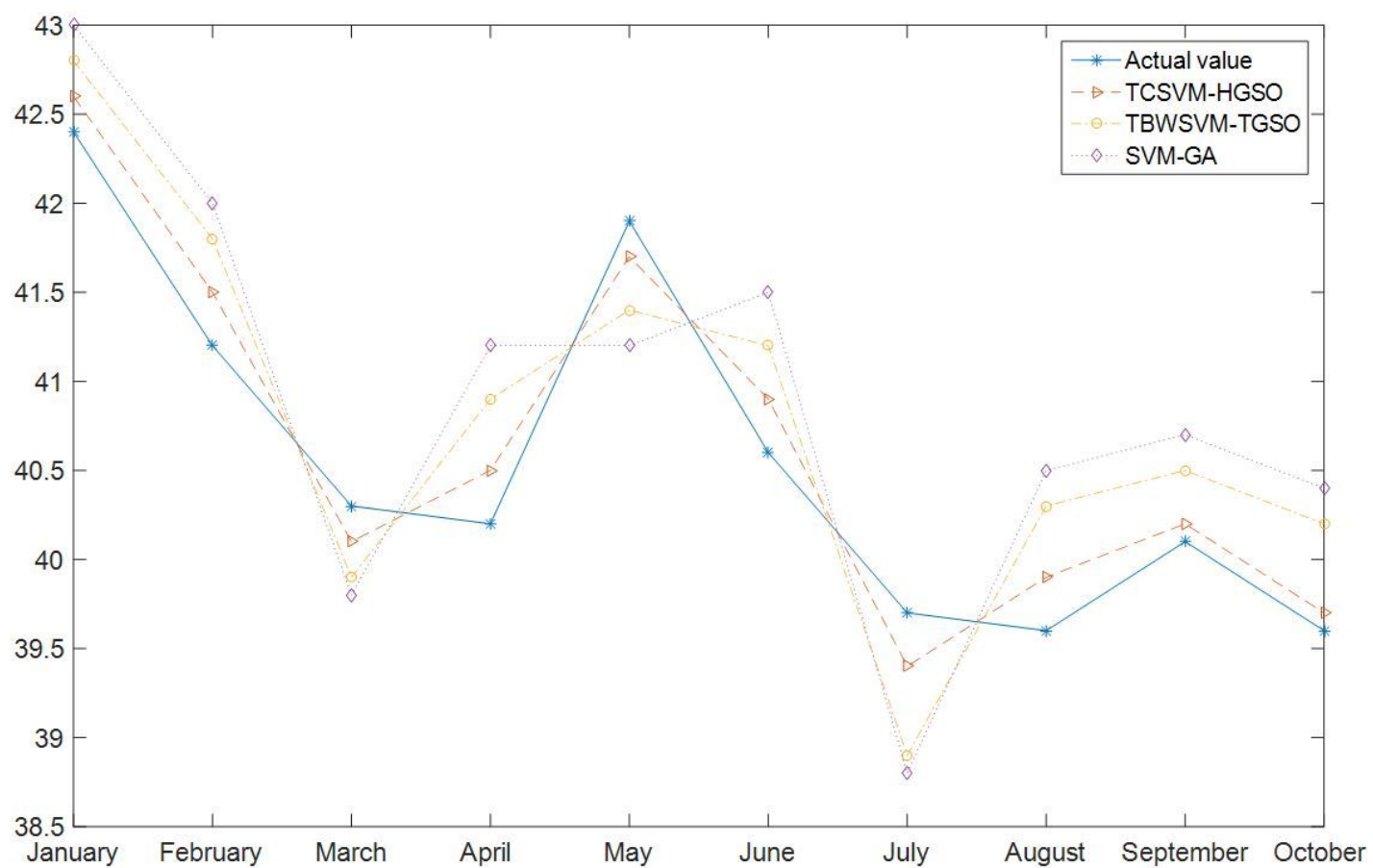


Figure 5

Prediction results of testing samples based on different models

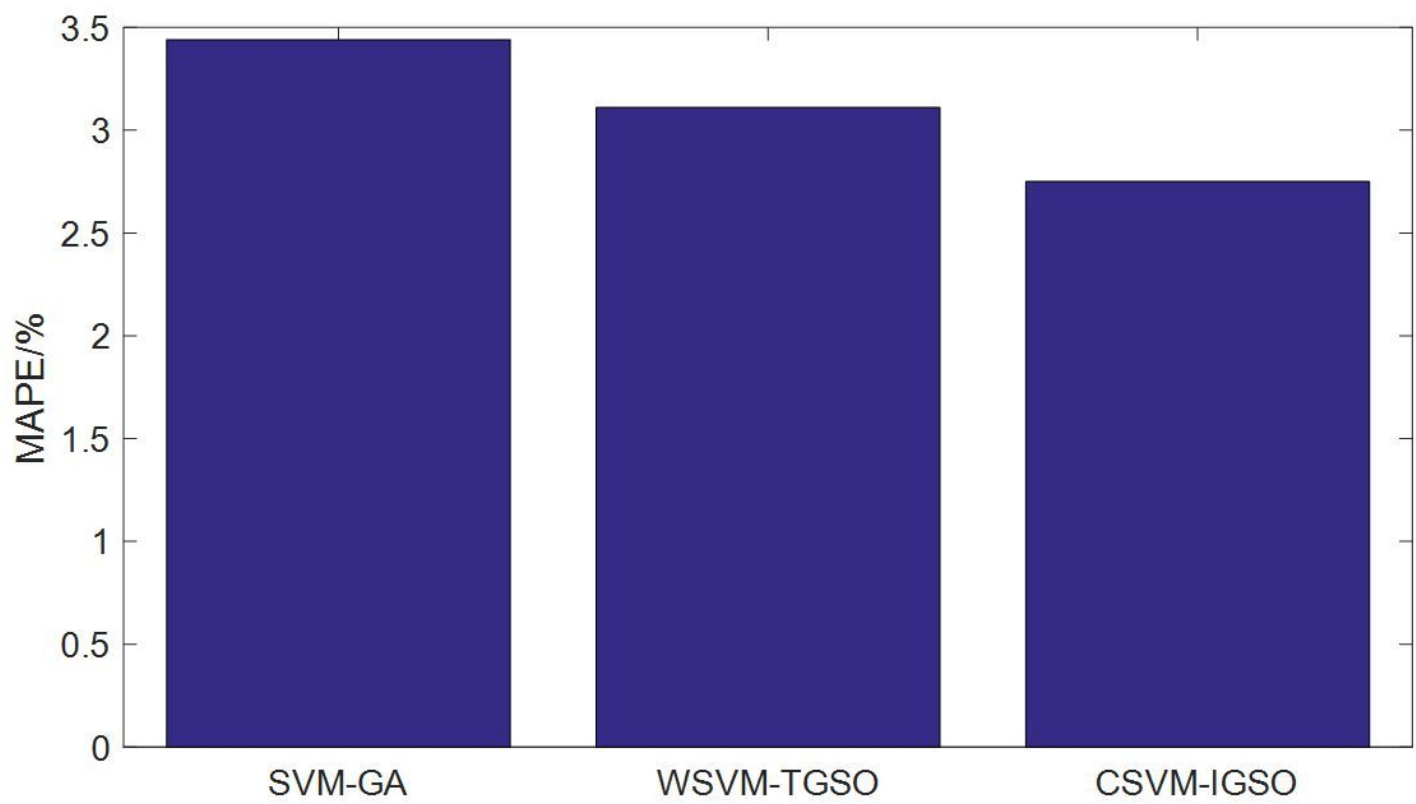


Figure 6

MAPE of different prediction models