Electronics and Electrical Engineering

AIMS Electronics and Electrical Engineering, 10(2): 265–284.
DOI: 10.3934/electreng.2026011
Received: 13 August 2025
Revised: 15 January 2026
Accepted: 29 January 2026
Published: 01 April 2026

https://www.aimspress.com/journal/ElectrEng

Research article

Early warning of transformer equipment based on robot infrared temperature prediction

Lijie Sun[1,*], Junfei Zhu[2], Weishang Gao[3], Qin Gao[3] and Márk Domonkos[4]

[1] School of Art and Design, Taizhou University, Taizhou, 318000, China
[2] State Grid Taizhou Power Supply Company, Taizhou 318000, Zhejiang, China
[3] School of Intelligent Manufacturing, Taizhou University, Taizhou 318000, China
[4] Robotics Research Center, Eötvös Loránd university, Budapest 1088, Hungary

* **Correspondence:** Email: lijiesun@tzc.edu.cn; Tel: +1-575-763-8496; Fax: +1-575-763-8496.

Abstract: Smart grid construction is a goal of the State Grid Corporation of China (SGCC). Transformer equipment, as the foundation of substation power supply and a critical component of the national power grid, requires stable operation. This paper proposes an early warning method for transformer equipment based on robot infrared temperature prediction, which mainly includes three parts: robot infrared temperature measurement, GWO-SVR temperature prediction based on historical equipment data, and determination of defect types according to heating defect classification standards for transformer equipment. This paper leverages the strong generalization ability of support vector regression (SVR), which maintains good prediction performance on small-sample data. To address limitations in the traditional SVR model caused by penalty factors and kernel function parameters, the Greey Wolf Optimizer (GWO) is introduced to optimize parameters and achieve optimal selection. The resulting GWO-SVR model is used for equipment temperature prediction, and defect types are identified in advance using dual-level grading standards from the state and the Taizhou Power Bureau, enabling early warning of transformer equipment. Compared with the GRNN, BP neural network, SVR model, LSTM, and Attention-LSTM, the GWO-SVR achieves better performance in substation temperature prediction, as evaluated by mean absolute error (MAE) and mean square error (MSE) metrics. In addition, GWO-SVR enables heat fault early warning based on temperature prediction and classification standards of heating defects for transformer equipment. This paper contributes to the unique application context of State Grid intelligence with previously unaddressed challenges.

Keywords: transformer equipment; temperature prediction; SVR; GWO; time series

1. Introduction

Nowadays, the state comprehensively promotes the construction of smart grids and the intelligent construction of power grid equipment, extending applications of new technologies, such as intelligent operation, monitoring, and inspection, strengthening the tackling of key core technologies, and accelerating the construction of smart substations. As the foundation of substation power supply, transformer equipment is a critical part of the national grid supply. Stable operation is critical because equipment status can be timely detected and abnormalities promptly eliminated in power system management and transformation [1]. However, power failures caused by equipment temperature overload often disrupt substation operation and may cause outages in nearby areas. In severe cases, fires and safety accidents can occur, leading to economic losses and social problems [2,3].

Currently, the State Grid adopts an intelligent inspection system to maintain the smooth operation of transformer equipment, in which the intelligent visual temperature measurement expert diagnosis system is a key component. In the early stage of intelligent construction (2019), the State Grid was in the pilot operation stage. As it is impossible to determine which monitoring method is more effective, a combination of fixed infrared camera monitoring and robot infrared temperature measurement was applied for intelligent infrared temperature inspection of substation equipment. The intelligent visual temperature measurement expert diagnosis system for main transformer equipment in a substation located in Taizhou City, Zhejiang Province, is shown in Figure 1, detailing equipment counts, monitoring methods, substation equipment, infrared imaging diagram, and fault categories.

Robot infrared temperature measurement is recognized as a future direction for intelligent grid construction. Real-time monitoring and temperature-based early warning can reduce equipment failures. However, short storage space and operational runtime of monitoring devices constrain timely fault diagnosis from infrared early-warning data. Meanwhile, the establishment and optimization of temperature monitoring and a fault early-warning management system remain essential for intelligent and information-driven substation management.

Currently, the temperature monitoring and prediction system in operation still has significant shortcomings. It cannot function independently of manual judgment or achieve real intelligence, especially in terms of fault early warning. Most substations still rely on specialists to review and interpret the data, resulting in high labor costs. Moreover, manual detection makes it difficult to effectively control the error rate of fault early warnings. Therefore, leveraging historical temperature data to find the relationship between temperature and equipment failure, and to improve the accuracy of future temperature predictions, would help operators to assess the operational status of electrical equipment in advance. Based on transformer fault classification categories and predefined temperature thresholds, early prediction and warning of potential fault types can reduce the likelihood of power supply safety incidents and better ensure normal substation operation [4,5].

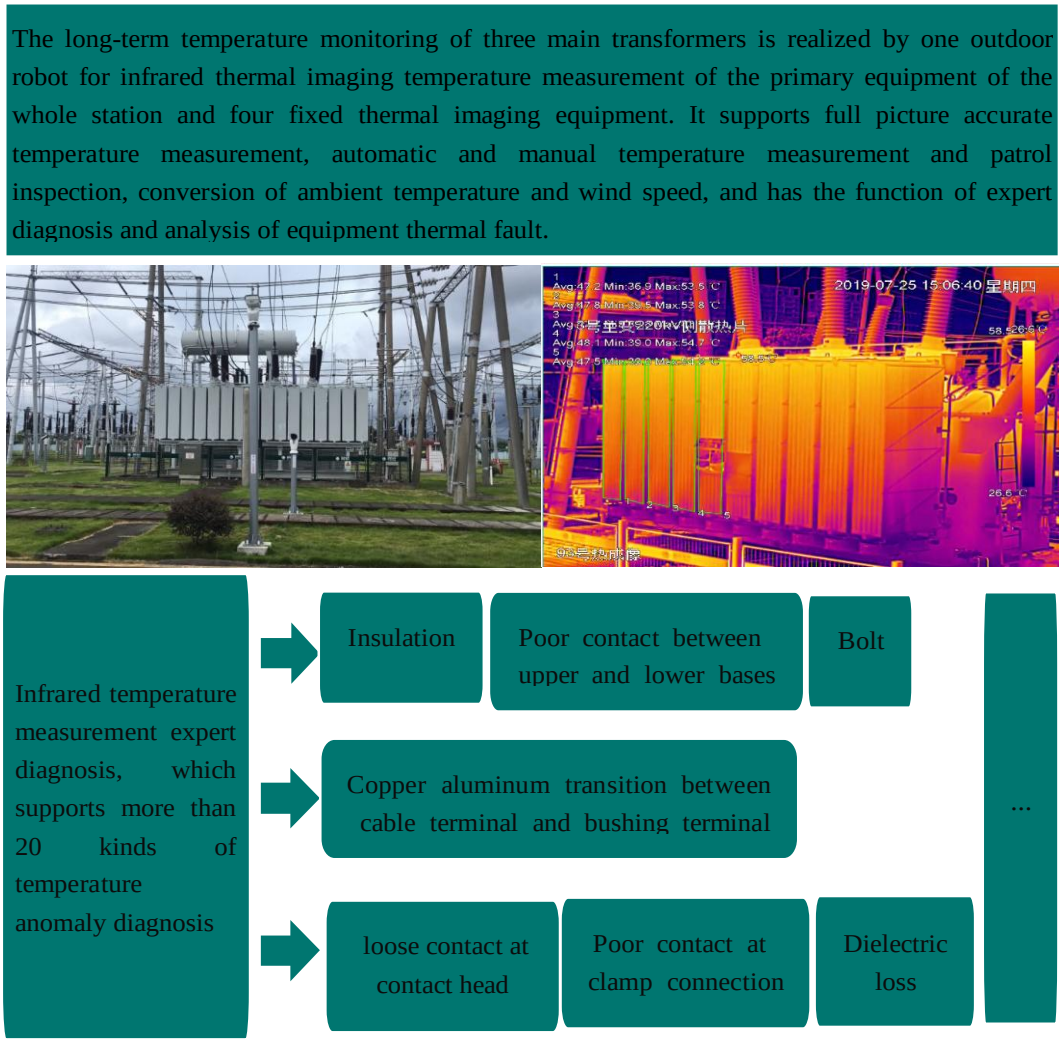


Figure 1. Intelligent visual temperature measurement expert diagnosis system.

In summary, from the perspective of accurate transformer equipment temperature prediction, and based on transformer fault classification principles and heating defect characterization, this study takes the primary main transformer equipment of a substation as the research object and proposes an infrared temperature–based early warning method. The work is developed from two aspects: (1) a temperature prediction model for substation equipment, and (2) threshold classification standards for equipment heating defects.

Although we are now in the era of big data, equipment condition monitoring in substations remains at an early stage of reform and development, and datasets are still limited. Consequently, it is particularly important to seek a model with strong generalization ability and good prediction performance under small sample data. Support vector regression (SVR) avoids local optimization issues and demonstrates strong generalization performance in small-sample, nonlinear, and high-dimensional regression problems. However, its performance is highly sensitive to parameter selection [6,7].

In recent years, many studies have focused on improving SVR prediction performance through optimization. In 2023, Chinese scholar Yunan Zheng et al. [8] used an SVR model to forecast natural gas prices at the Henry Hub during the Russia–Ukraine conflict, providing valuable theoretical references for global energy policymakers and natural gas market investors. In 2024, Chinese scholar

Xianwei Wu et al. [9] used an SVR prediction model to predict flow-induced vibration in a steam generator heat transfer tubes caused by high-temperature and high-pressure fluids; the average relative error of the prediction model was only 2.1%, demonstrating high accuracy, which is helpful for safety early warning and maintenance of steam generator tubes in nuclear power plants. In the same year, Chinese scholar Wencong Wu et al. [10] employed SVR to predict the mixing time and degree of binary spherical mixtures in a horizontal rotating drum in a stable mixing state, obtaining good performance. In 2025, Iranian scholars applied an SVR model to predict sharp-edged width constrictions, highlighting its effectiveness in accurately predicting Cd and its strong generalization capabilities in hydraulic applications [11].

The grey wolf optimizer (GWO) features a relatively simple structure, requires fewer parameters than other optimization algorithms, and has good robustness, making it straightforward to implement in repeated experiments. Compared with other algorithms [12], such as particle swarm optimization (PSO) [13], genetic algorithm (GA) [14], and firefly algorithm [15], GWO integrates the three best optimization schemes for optimization, thereby enhancing global optimization ability. Due to these advantages, GWO is employed in this study to determine the key SVR factors, namely the penalty factor and kernel function parameter.

Substation equipment temperature prediction is a typical time-series prediction problem. Traditional methods, such as the auto-regressive and moving average (ARMA) model, are commonly used [16], but such single-factor methods rely solely on the historical data and cannot fully characterize the complexity of the data. Based on previous studies [17–20] and by analyzing experimental data, an obvious seasonal variation trend is observed. Therefore, environmental factors are incorporated into the prediction framework [21]. Specifically, daily maximum and minimum temperatures are combined with historical equipment temperature to construct the input eigenvector, and a GWO-SVR model is adopted for temperature prediction [22]. Prediction experiments conducted on phase A, phase B, and phase C contactors demonstrate the prediction advantages of the GWO-SVR model.

In November 2008, the National Development and Reform Commission of the People's Republic of China published the standard DL/t664-2008, *Application Specification for Infrared Diagnosis of Live Equipment*, which provides classification standards for heating defects of transformer equipment in Appendix A [23]. In addition, using the transformer equipment of the Taizhou substation as the research object, heating defect thresholds for different components can be clearly defined in accordance with the defect grading regulations of the Taizhou Power Bureau. Accordingly, this study predicts temperature trends based on the GWO-SVR temperature prediction model and implements transformer equipment early warning according to the two-level heating defect grading standards established by the state and the Taizhou Power Bureau.

2. Research methods

2.1. SVR

The SVR model is a supervised machine learning method for solving function regression problems [24,25]. The implementation process is as follows:

Define the sample set as $S = \left\{ \left(x_i, y_i \right) \middle| x_i \in R^n, y_i \in R \right\}_{i=1}^{L}$, where the input eigenvalue of the $i-th$ sample is denoted as x_i, and the real value of the $i-th$ sample is denoted as y_i; the dimension of the

input parameter is n, and the training sample number is L. The eigenvector x of each sample in the original space is mapped to the high-dimensional space, and the expression of the linear regression function of the support vector machine in a high-dimensional feature space is

$$f(x) = \omega\varphi(x) + b \tag{1}$$

where the weight vector is expressed by w, the offset is expressed by b, and the nonlinear mapping function is expressed by φ. $\varphi(x)$ is recorded as the mapped eigenvector.

The minimization objective function used in solving ω and b is denoted as formula (2):

$$\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{L}\left| f(x_i) - y_i \right|_{\varepsilon} \tag{2}$$

where the maximum error supported by the regression analysis is represented by ε; C represents the penalty coefficient; $f(x_i)$ indicates the predicted value of the i-th training sample; and $\left| f(x_i) - y_i \right|_{\varepsilon}$ is the constraint condition, which has the function of maintaining good sparsity in SVR models. When the constraint condition is not met, introduce the relaxation variables ξ_i and ξ_i^*, and then formula (2) is converted to formula (3).

$$\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{L}\left| \xi_i - \xi_i^* \right|_{\varepsilon} \tag{3}$$

ξ_i, ξ_i^*, $f(x_i)$, y_i, and ε meet the following conditions shown in formula (4):

$$\text{s.t.}\begin{cases} y_i - f(x_i) \le \varepsilon + \xi_i \\ -y_i + f(x_i) \le \varepsilon + \xi_i^* \quad i = 1,2,...,L \\ \xi_i,\ \xi_i^* \ge 0 \end{cases} \tag{4}$$

The kernel function is introduced through the duality principle and Lagrange function, and the input vectors x_i and x_j are mapped from the original space to higher dimensions using the implicit function φ. $\varphi(x_i)$ and $\varphi(x_j)$ are denoted as the mapped feature vectors, and $K(x_i, x_j) = \varphi(x_i)^T \bullet \varphi(x_j)$ is the kernel function. Due to the excellent nonlinear approximation and generalization ability of the RBF function in parameter selection, this paper adopts RBF as the kernel function, as shown in formula (5):

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2} \right) \tag{5}$$

where σ is the kernel function parameter.

In the SVR model, the size of the penalty factor C indicates the punishment degree when the error exceeds ε. When C is very large, the prediction accuracy is not enough, and when C is very small, the generalization ability of the model is weak. Meanwhile, when σ in the kernel function is

very large, the mapping result is not ideal, which is actually equivalent to a low-dimensional space; when σ is very small, there may be fitting problems. Therefore, how to choose the appropriate penalty factor and kernel function parameters is an important factor in increasing the model's performance.

2.2. GWO

The grey wolf optimizer (GWO) is a swarm intelligence optimization algorithm proposed by Mirjalili et al. in Australia in 2014 [26], which has been successfully applied in various fields to solve parameter optimization problems [27]. GWO realizes the hunting task by simulating the leadership and hunting mechanism of grey wolves in nature, strictly observing the hierarchy of social dominance; parameter optimization is achieved based on this principle. The GWO optimization process consists of four steps: social hierarchy, surrounding, hunting, and attacking prey.

(1) Social hierarchy

To apply the GWO, first, it is necessary to establish a hierarchical model of social dominance of grey wolves, as shown in Figure 2. Fitness is calculated for each individual in the population, and the three grey wolves with the best fitness in the wolf pack are marked as α, β, δ; the rest are marked as ω. The social rank of the grey wolf group is as follows, from high to low: $\alpha, \beta, \delta, \omega$. The optimization process is mainly guided by α, β, and δ.

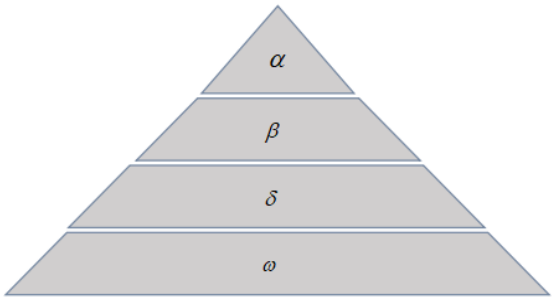


Figure 2. Grey wolf social dominance hierarchy diagram.

(2) Surrounding prey

Grey wolves can be close to their prey gradually and surround it when they tie it with ropes. This behavior is described by a mathematical model, which is expressed by formulas (6)–(9).

$$D = C \cdot X_p(t) - X(t) \tag{6}$$

$$X(t+1) = X_p(t) - A \cdot D \tag{7}$$

$$A = 2a \cdot r_1 - a \tag{8}$$

$$C = 2r_2 \tag{9}$$

where t is the iteration sequence number; $\cdot$ is the product operation of Hadamard; X_p represents the vector of prey position information; $X(t)$ represents the vector of the current grey wolf position; A and C are the vectors of synergy coefficient; a decreases linearly during the whole iteration

process, in which the range is $[2,0]$; and r_1 and r_2 are the vectors with a random range $[0,1]$.

(3) Hunting prey

The grey wolf possesses the ability to search for the location of potential prey (optimal solution). The process is mainly finished by the guidance of grey wolves α, β, δ, but the solution space characteristics are unknown for many problems, and the grey wolf cannot determine the exact location of prey. In order to simulate the search behavior (candidate solution), it is assumed that α, β, and δ have a strong ability to identify the location of potential prey. Therefore, in each iteration, the best three grey wolves (α, β, δ) in the current population are retained, and then the locations of other search agents (including ω) are updated according to their location information. The expression of the mathematical model for this behavior is shown from formula (10) to formula (16.

$$D_\alpha = |C_1 \cdot X_\alpha(t) - X(t)| \tag{10}$$

$$D_\beta = \left|C_2 \cdot X_\beta(t) - X(t)\right| \tag{11}$$

$$D_\delta = |C_3 \cdot X_\delta(t) - X(t)| \tag{12}$$

$$X_1 = X_\alpha - A_1 \cdot D_\alpha \tag{13}$$

$$X_2 = X_\beta - A_2 \cdot D_\beta \tag{14}$$

$$X_3 = X_\delta - A_3 \cdot D_\delta \tag{15}$$

$$X_P(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{16}$$

where $X_\alpha, X_\beta, X_\delta$ represent the vectors of position information for α, β, δ in the current population, respectively; X represents the vector of position information for grey wolf; C_1, C_2, C_3 are random vectors; and $D_\alpha, D_\beta, D_\delta$ represent the distance information between the current candidate grey wolf and the best three wolves, respectively. When $|A| > 1$, grey wolves are scattered in various areas as far as possible and search for prey; when $|A| < 1$, grey wolves concentrate on hunting prey in one or some areas; $X_P(t+1)$ is the updated position vector of the individual. The grey wolf hunting process is shown in Figure 3.

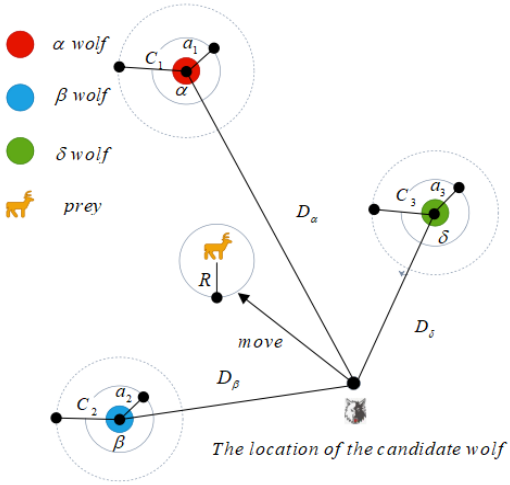


Figure 3. Hunting prey process.

It is obvious that the position of the candidate solution finally falls in the position of the random circle defined by α, β, and δ from Figure 3. In general, α, β, and δ need to first predict the approximate location of the prey (potential optimal solution), and then other candidate wolves randomly update their location near the prey under the guidance of the current optimal three wolves.

(1) Attacking prey

In the process of building the attack prey model, according to formula (2), the decrease of value a will cause the fluctuation of value A. In other words, A is a random vector with the range $[-a, a]$, where a decreases linearly during the iteration. When A belongs to $[-1,1]$, the next position of the search agent can be anywhere between the current grey wolf and the prey.

(2) Search for prey

Grey wolves mainly rely on information from A, B, and C to find prey. They begin to search for prey location first in a scattered way and then concentrated on attacking prey. For the establishment of the decentralized model, $|A|>1$ is used to keep its search agent away from prey, which enables GWO to conduct a global search. Another search coefficient in the GWO algorithm is C. It can be seen from formula (2) that the C vector is composed of random values in the interval range $[1,3]$; this coefficient provides a random weight for prey to increase ($|C|>1$) or decrease ($|C|<1$), which helps GWO show random search behavior in the optimization process to avoid the algorithm falling into local optimization. It is worth noting that C does not decrease linearly. C is a random value during the iteration process. This coefficient is conducive to the algorithm to jump out of the local, especially in the later stages of the iteration.

3. GWO-SVR prediction method

In the prediction of GWO-SVR, GWO is adopted to optimize key parameters of C and σ in SVR. The GWO-SVR model is constructed and used for equipment temperature prediction. The prediction implementation process of GWO-SVR is as shown in Figure 4; the implementation process includes the following six steps:

(1) Import and preprocess the data.
(2) Initialize the information of wolves, including the number and position, the maximum number of iterations, and the upper and lower bounds of C and σ.
(3) Calculate the fitness of each grey wolf according to the mean square error (MSE) and select the top three grey wolves with the best fitness.
(4) Calculate D_α, D_β, D_δ, the orientation of other grey wolves, and the parameters α, A, C according to formulas (9)–(15).
(5) Judge whether the maximum number of iterations has been reached. If so, save the current value as the optimal solution of C and σ; otherwise, execute step 3.
(6) Take C and σ under the optimal solution to conduct optimal GWO-SVR prediction model.

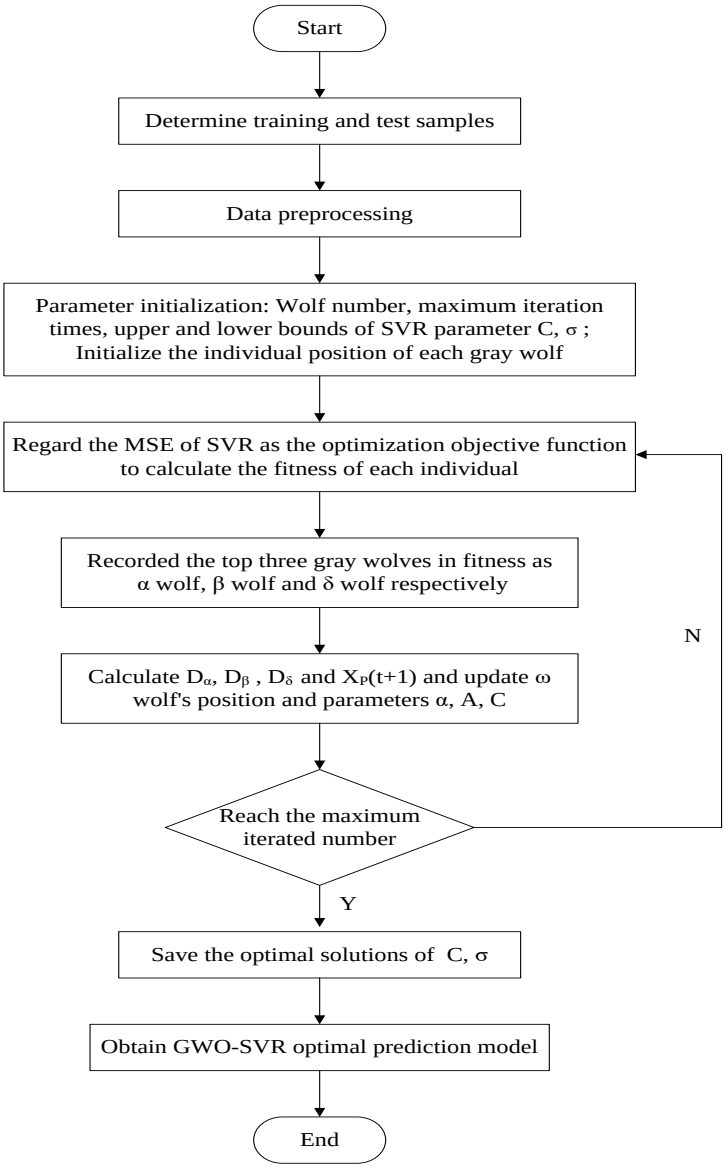


Figure 4. GWO-SVR prediction flow chart.

4. Temperature prediction experiments

4.1. Data sources and analysis

This study focuses on the primary equipment of the main transformer in a substation. The dataset consists of one year of temperature measurements from the A, B, and C-phase contactors of the equipment bushings in Taizhou City, Zhejiang Province. Figure 5 shows the temperature variation of the phase C contactor.

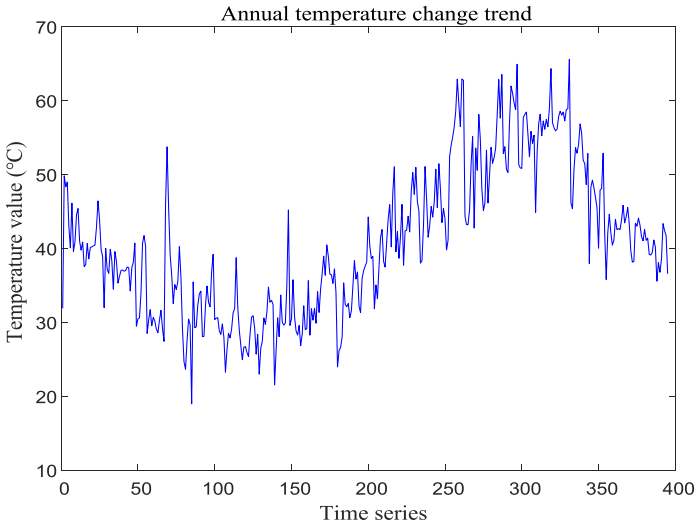


Figure 5. Temperature variation of the bushing phase C contactor.

The same measuring point exhibits different temperature levels across seasons. The specific temperature change is shown in Figure 6, where blue represents summer and red represents winter. The figure indicates that equipment temperature varies with the season, demonstrating a relationship between equipment temperature and ambient temperature. Therefore, ambient temperature is a critical factor in equipment temperature prediction.

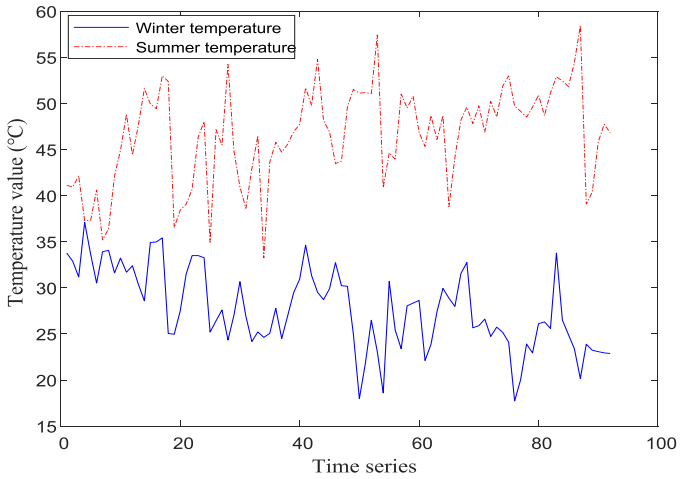


Figure 6. Temperature variation of the phase C contactor of a 220 kV equipment bushing in summer and winter.

The data collected by substations is not only large but also heterogeneous, including equipment operation data, basic equipment data, and defect records, which are typically stored in different systems. In general, temperature data is relatively stable. Traditional methods assume that there is a linear relationship within the data and apply regression-based fitting methods. However, actual temperature data is influenced by the equipment itself and ambient temperature. As a result, its internal change is irregular and characterized by a nonlinear relationship.

4.2. Data process

4.2.1. Selection of prediction samples

According to the above analysis, several factors affect the substation equipment temperature, among which the ambient temperature and the equipment temperature have an important impact on the future equipment temperature. Therefore, this paper selects the daily maximum and minimum temperatures, as well as the equipment temperature in the past three days, to form an eigenvector with five feature parameters for temperature prediction.

Data was collected for substation equipment in Taizhou City, Zhejiang Province, from October 1, 2019, to October 29, 2020. There are 389 data points from phase A contactor, phase B contactor, and phase C contactor, from which 327 are used for training the model, and the remaining are used as prediction samples. In addition, ambient temperature data (daily maximum temperature and minimum temperature) from October 1, 2019, to October 29, 2020, were collected. The same data proportion was used to divide training and testing samples, form the eigenvector, and predict the substation equipment temperature.

4.2.2. Data standardization

The feature vector used for substation equipment temperature prediction includes five input features. However, because each feature has different units and is limited, the data needs to be standardized. The most typical normalization processing method [28,29] was used to scale the data to make it fit the interval of $[0,1]$ and become a dimensionless value, and then input it into the model for prediction. In this paper, min-max is adopted; the standardized formula is shown in formula (17):

$$X^* = \frac{X - min}{max - min} \tag{17}$$

where X represents the original data, $\min$ represents the minimum of X, and $\max$ represents the maximum of X.

4.3. Prediction performance evaluation index

Mean absolute error (MAE) and mean square error (MSE) are used to verify the prediction effect of the models [30,31]. MAE and MSE are calculated as shown in formulas (18) and (19).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \tag{18}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{19}$$

In the above formulas, N is the predictive sample number, and y_i and $\hat{y}_i$ refer to the true value and predictive value of $i-th$ sample, respectively.

4.4. Comparative experiments

This paper compares GWO-SVR with the BP neural network (BPNN), generalized regression neural network (GRNN), and the traditional SVR model, as well as long short-term memory (LSTM) and LSTM with attention mechanism (Attention-LSTM) [32], under the same parameter initialization conditions:

(1) The number of iterations of all models is 1000.

(2) According to preliminary experiments, the GWO initialization range in GWO-SVR is $[5,10]$.

(3) The grid search is adopted to search SVR parameters, which include C and σ, and the search range is $[-5, 0.5, 5]$.

(4) The prediction performance of BPNN is based on the average of ten prediction results, in which the learning rate is 0.2, the optimal network structure is determined as 5-7-1, and the temperature is predicted.

(5) The GRNN model adopts k-fold cross-validation, in which $k = 3$. In addition, the input layer is 5, the output layer is 1, and the key parameter for controlling the response range of neurons is spread from 0.1 to 2, with a diffusion interval of 0.1.

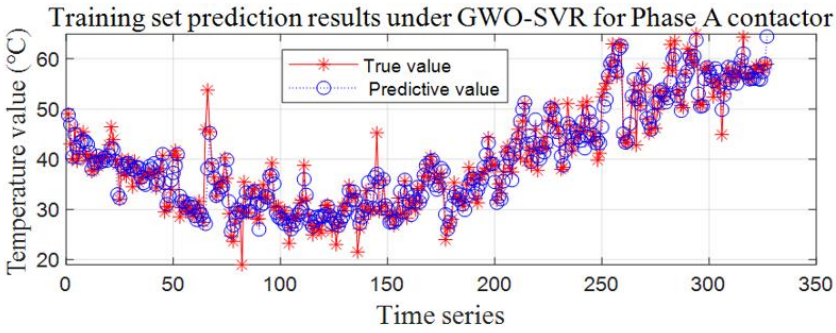
(6) For the LSTM model, the training cycle is 20 times per round, for a total of 50 rounds, with a total of 1000 iterations. The initial learning rate of the algorithm is set to 0.005, the input is 5, the output is 32, and a fully connected layer is added to obtain the temperature prediction value. In addition, the parameters of Attention-LSTM are the same as those of the LSTM model.

The training and testing prediction results of GWO-SVR from 3-phase contactors are shown in Figures 7–9. Key parameters of the six methods are listed in Table 1; the prediction performance of the six methods is shown in Table 2. Three sets of phase contactor data were collected from phase A, phase B, and phase C contactors. Evaluating indicators adopted MSE and MAE according to formulas (18) and (19). Running time was counted in seconds, with the mean of running time.

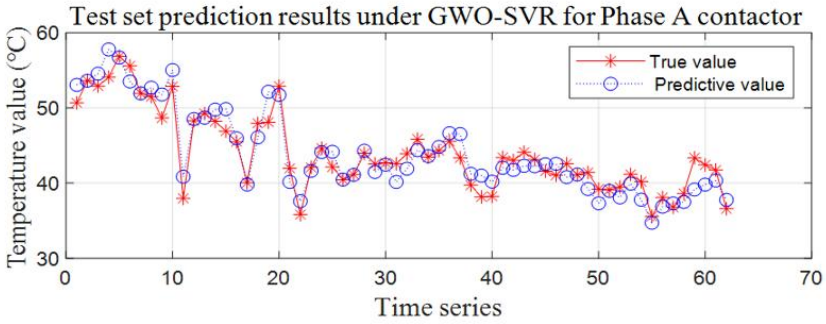
Table 1. Key parameters for the six methods.

Method	Key parameters	Value
GWO	Initialization range	$[5,10]$
SVR	Iteration count	1000
	Penalty factor C	search range: $[-5, 0.5, 5]$
	Kernel function	RBF function
	Kernel function parameter σ	search range: $[-5, 0.5, 5]$
	Maximum error ε	0.1
	Learning rate	0.2
	Optimal network structure	5-7-1

BPNN	Training goal (max MSE)	0.00005
	Iteration count	1000
GRNN	Radial basis function spread range	0.1-2
	Radial basis function spread step size	0.1
	k-fold cross-validation, k	3
	Network structure	input:5,output:1
	Training goal (max MSE)	0.00005
LSTM	Iteration count	1000
	Network structure	5-128-32-1
	Initial learning rate	0.005
	Training cycle time count per round	20
	Total round count	50
	Optimizer	Adam
Attention-LSTM	Iteration count	1000
	Network structure	5-128-32-1
	Initial learning rate	0.005
	Training cycle time count per round	20
	Total round count	50
	Optimizer	Adam
	Attention mechanism	Points system
	Training goal (max MSE)	0.00005

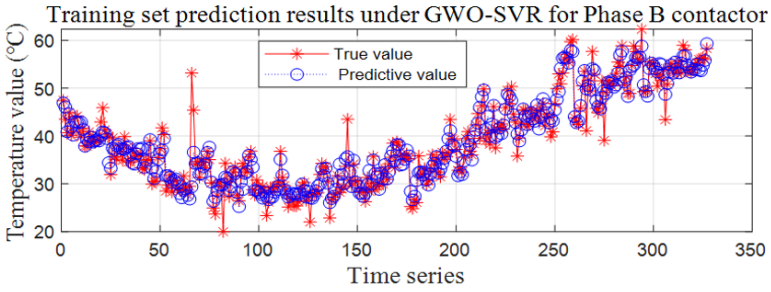


The fitting results of the training set in Phase A contactor

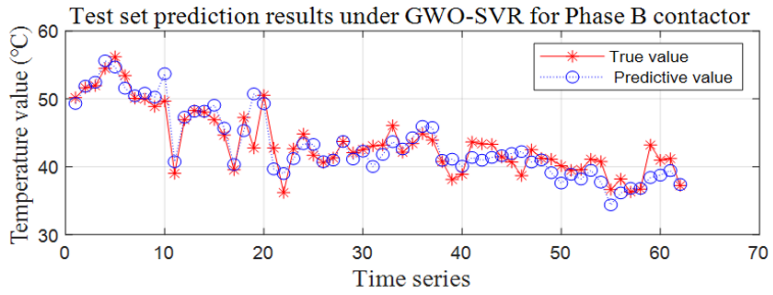


The prediction results of the test set in Phase A contactor

Figure 7. Prediction results of GWO-SVR for the phase A contactor.

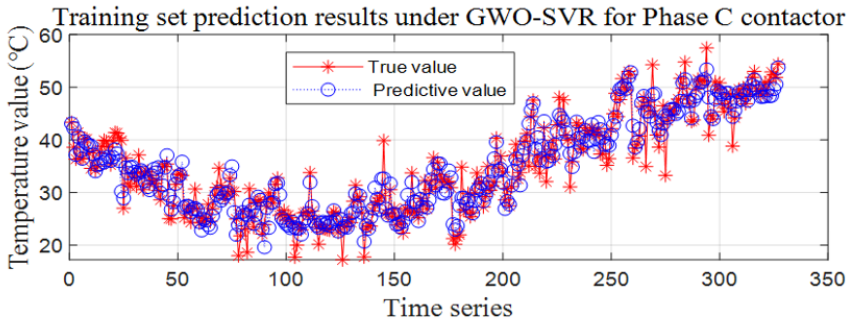


The fitting results of the training set in Phase B contactor

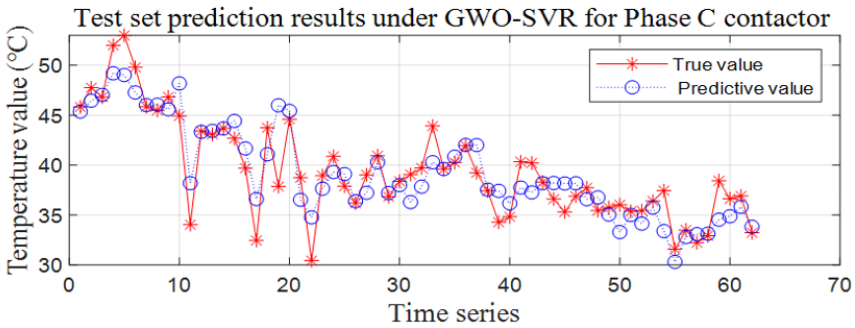


The prediction results of the test set in Phase B contactor

Figure 8. Prediction results of GWO-SVR for the phase B contactor.



The fitting results of the training set in Phase C contactor



The prediction results of the test set in Phase C contactor

Figure 9. Prediction results of GWO-SVR for the phase C contactor.

Table 2. Prediction performance comparison of the six methods for the three sets of phase contactor data.

Data set	Evaluating indicator	BPNN	GRNN	LSTM	Attention-LSTM	SVR	GWO-SVR
Phase A contactor	MSE	9.7162	5.2102	17.3375	9.9994	5.0995	3.1724
	MAE	2.1143	1.7725	2.5533	2.2003	1.6865	1.4487
	Running time (s)	0.52371	4.9262	11.350324	17.038724	6.7925	2.9281
Phase B contactor	MSE	19.3714	7.5241	13.1814	7.7220	4.2001	4.1200
	MAE	2.4690	2.0481	2.2316	2.0458	1.5512	1.5286
	Running time (s)	0.7557	5.0009	10.533335	17.062008	6.304	3.0514
Phase C contactor	MSE	13.9838	6.8685	15.9551	8.8931	5.1257	5.0962
	MAE	2.6222	1.9409	2.5511	2.1324	1.6940	1.6845
	Running time (s)	0.50578	5.1379	10.134678	16.947403	7.3078	3.2008

According to the experimental analysis for the three sets of phase contactor data, the following results were obtained:

(1) The SVR model optimized with GWO has improved prediction performance and faster search speed compared to SVR using a traditional grid search.
(2) The selection of the BP neural network weight has obvious randomness, which leads to instability of the prediction effect.
(3) Compared with other models, GWO-SVR is the best model for substation temperature prediction.
(4) LSTM is a type of time-recurrent neural network specifically designed to solve the long-term dependency problem of general recurrent neural networks (RNNs). It is mainly used for time series prediction and has applications in stock price prediction, weather prediction, and traffic flow prediction, among others. However, for the temperature prediction problem of the three experiments, the running time of the traditional BPNN was approximately 0.5 s, while the LSTM required approximately 10 s, showing no improvement compared with traditional neural network models; due to its complex network structure, time complexity increased by 20 times.
(5) Because the substation thermal imaging monitoring equipment has only been in use for close to one year, there is a small amount of data, resulting in a small number of samples used for training; this leads to somewhat poor prediction results. Nevertheless, the performance of SVR and GWO-SVR is relatively good, while the GWO-SVR shows the best performance.

5. Transformer equipment early warning

The early-warning process for transformer equipment heating defects based on the GWO-SVR temperature prediction primarily includes infrared temperature measurement of equipment using robots, GWO-SVR temperature prediction based on historical equipment temperature, and fault early warning according to the evaluation criteria for transformer heating defects. Section 4 verified the effectiveness of the GWO-SVR method in predicting transformer equipment temperature. The defect evaluation principles are described below.

5.1. Defect classification principles for transformer equipment

Any abnormal condition occurring in operating substation equipment that affects safe operation, regardless of whether it can continue operating, is defined as an equipment defect. According to severity, defects can be classified into three categories:

(1) Emergency defects: Defects that pose a serious threat to personnel, the power grid, or equipment, that can cause accidents at any time, and that must be handled immediately.
(2) Important defects: Defects that have a certain impact on equipment service life and safety, reduce the output of the equipment, or may develop into a threat to personnel, the power grid, or the equipment. Although these allow maintenance of operation for a period of time, they should be handled as soon as possible.
(3) General defects: Defects that pose minimal risk to equipment operation or safety and are unlikely to develop into important defects in the short term. These can be addressed during equipment maintenance and testing.

5.2. Early warning based on GWO-SVR prediction

Based on the two-level heating defect grading standard established by the state and the Taizhou Power Bureau, and in accordance with the infrared diagnosis guidelines for live equipment [33], this study determines heating defect types using three indicators: equipment surface, maximum temperature rise, and interphase temperature difference. The specific description of some heating defects is shown in Table 3.

Table 3. Description of some heating defects of transformer equipment.

Transformer equipment components	Emergency defect	Important defect	General defect
Enclosed switchgear	The temperature rise of parts easily accessible to operators exceeds 40 K	The temperature rise of parts easily accessible to operators exceeds 30 K	The temperature rise of parts easily accessible to operators exceeds 20 K
Metal conductor; connection point	>110 ℃	>80 ℃	The temperature difference between phases shall not exceed 15 K
Electrical connection point of heat-resistant conductor; connectors and clamps	>130 ℃	>90 ℃	The temperature difference between phases shall not exceed 15 K
Bushing	>80 ℃	>55 ℃	The temperature difference between phases shall not exceed 10 K

The experimental dataset consists of surface temperature measurements collected from the bushing positions of primary substation equipment. The temperature trend of the bushing contactors is predicted using the GWO-SVR model, and the heating defect type can be determined according to Table 3.

Transformer bushing contactors belong to the connection point category. According to Table 2, when the temperature is higher than 110 ℃, it is an emergency defect; when the temperature is higher than 80 ℃, it is an important defect; and when the temperature difference between phases is higher

than 15 K, it is a general defect. At present, the intelligent robot infrared temperature detection deployed in the State Grid is still in a preliminary operation stage, and the availability of data is limited. Only individual values of the phase C contactor exceed 55 ℃ in summer, and no emergency defects have been observed. Taking the phase C contactor as an example, when the predicted result is within the safe range, the temperature warning result displays the message, "According to the prediction results, the phase C contactor temperature of the transformer will be within a safe range in the next day, and no maintenance instructions need to be sent to the site." This temperature warning is shown in Figure 10. When the predicted temperature is greater than 110 ℃, the temperature warning result displays, "According to the prediction results, the phase C contactor temperature of the transformer will be in danger in the next day, classified as an emergency defect, and maintenance instructions need to be sent to the site." When the predicted temperature is greater than 80 ℃, the temperature warning result displays, "According to the prediction results, the phase C contactor temperature of the transformer will be at high-level risk in the next day, classified as an important defect, and maintenance instructions need to be sent to the site."

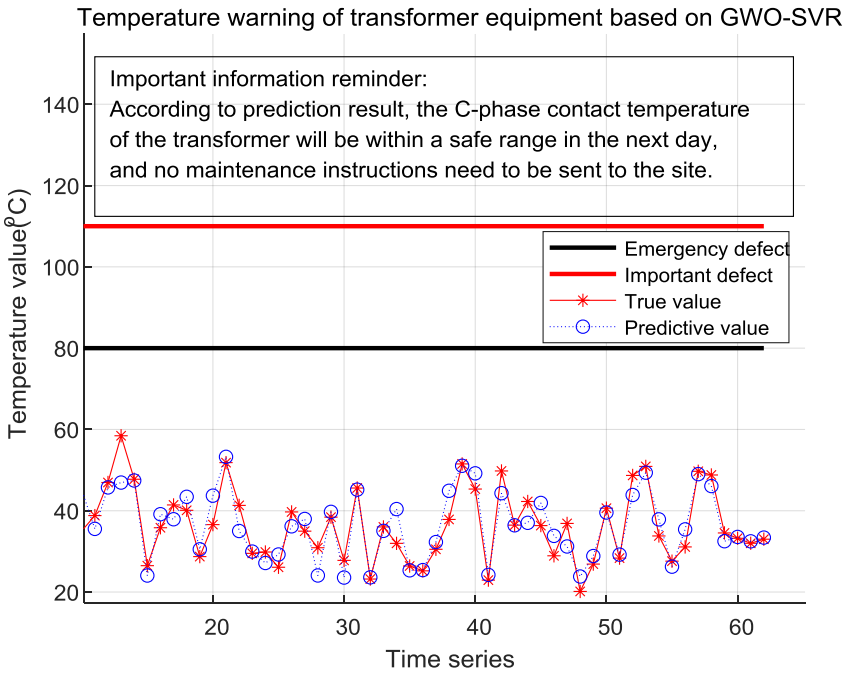


Figure 10. Temperature warning of transformer equipment based on GWO-SVR.

6. Conclusions

Transformer equipment failures can cause significant social consequences. As the State Grid advances intelligent grid construction, substations are gradually adopting intelligent inspection systems. This paper proposes an early warning method for transformer equipment based on robot infrared temperature prediction. The main contributions of this paper are summarized as follows:

Temperature prediction of substation equipment based on GWO-SVR: According to the complexity of substation equipment temperature data and the small-sample characteristics, SVR was selected for temperature prediction. To address the difficulty of selecting key parameters, GWO was used to obtain the optimal solution, forming the GWO-SVR method. Through temperature trend

prediction on three sets of phase contactor data, the method demonstrated better performance in terms of MAE, MSE, and running time.

By comparing with traditional neural network models and a deep learning network model (LSTM), the SVR model is more effective for small-sample time-series prediction in terms of prediction accuracy and time complexity. This result confirms that, although deep learning has many advantages, it is not suitable for some prediction problems and has limitations.

Based on the grading standards for heating defects established by the state and the Taizhou Power Bureau, and in accordance with the infrared diagnosis guidelines for live equipment, this paper adopts the equipment surface, maximum temperature rise, and interphase temperature difference to determine the heating defect type. Therefore, the proposed early warning framework for transformer equipment defects based on robot infrared temperature prediction provides reference values for intelligent state grid construction.

Author contributions

Lijie Sun: Writing - original draft, methodology, software, validation, experiment; Junfei Zhu: Visualization, investigation, writing-review & editing, software; Weishang Gao: Analysis, supervision, proofreading, project administration; Qin Gao: Editing and typesetting; Márk Domonkos: English polishing.

Use of Generative-AI tools declaration

The authors declare they have not used artificial intelligence (AI) tools in the creation of this article.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. Wu Y, Lin J, Liu J (2025) Switch monitoring algorithm for 220 kV terminal substation startup process based on multi time scale graph mode. *Electr Pow Syst Res* 239: 1–8. https://doi.org/10.1016/j.epsr.2024.111181
2. Sun S, Chen P, Lin Z, Zhang Z, Li Q (2025) Evolution characteristics and BLEVE risk analysis of overheating transformer oil under sudden depressurization. *Fuel* 393: 135075. https://doi.org/10.1016/j.fuel.2025.135075
3. Wu B, Wang X, Qi S, Ma X (2025) Emergency backup power robust planning for urban agglomeration power grids with a high proportion of new energy sources in extreme disaster scenarios. *Electr Pow Syst Res* 247: 111715. https://doi.org/10.1016/j.epsr.2025.111715
4. Amin A, Bibo A, Panyam M, Tallapragada P (2023) An Uncertainty-Aware Health Monitoring Model for Wind Turbine Drivetrains Based On Bayesian Neural Network. *IFAC PapersOnLine* 56: 235–240. https://doi.org/10.1016/j.ifacol.2023.12.030
5. Zhang X, Pan Y, Cao Y, Liu B, Yu X (2024) Cloud-based battery failure prediction and early warning using multi-source signals and machine learning. *J Energy Storage* 93: 112004. https://doi.org/10.1016/j.est.2024.112004

6. Sahoo SR, Gupta BB (2021) Multiple features based approach for automatic fake news detection on social networks using deep learning. *Appl Soft Comput* 100: 106983. https://doi.org/10.1016/j.asoc.2020.106983
7. Cvitić I, Peraković D, Periša M (2021) Ensemble machine learning approach for classification of IoT devices in smart home. *Int J Mach Learn Cyb* 12: 3179–3202. https://doi.org/10.1007/s13042-020-01241-0
8. Zheng Y, Luo J, Chen J, Chen Z, Shang P (2023) Natural gas spot price prediction research under the background of Russia-Ukraine conflict-based on FS-GA-SVR hybrid model. *J Environ Manage* 344: 118446. https://doi.org/10.1016/j.jenvman.2023.118446
9. Wu X, Xu Z, Liu Y, Chen Z, Huang J, Chen Y (2024) Failure prediction method of heat transfer tube of nuclear power steam generator based on WOA-SVR. *J Radiat Res Appl Sci* 17: 100907. https://doi.org/10.1016/j.jrras.2024.100907
10. Wu W, Chen K, Tsotsas E (2024) Prediction of particle mixing in rotary drums by a DEM data-driven PSO-SVR model. *Powder Technol* 434: 119365. https://doi.org/10.1016/j.powtec.2024.119365
11. Afaridegan E, Fatahi-Alkouhi R, Azarm P, Amanian N (2025) Enhanced prediction of discharge coefficient in sharp-edged width constrictions using a novel hybrid SVR-IWOA and machine learning models. *J Hydrol* 657: 133103. https://doi.org/10.1016/j.jhydrol.2025.133103
12. Judge MA, Franzitta V, Curto D, Guercio A, Cirrincione G, Khattak HA (2024) A comprehensive review of artificial intelligence approaches for smart grid integration and optimization. *Energy Conversion and Management: X* 24: 100724. https://doi.org/10.1016/j.ecmx.2024.100724
13. Ma H, Tang JM (2020) Short-Term Load Forecasting of Microgrid Based on Chaotic Particle Swarm Optimization. *Procedia Computer Science* 166: 546–550. https://doi.org/10.1016/j.procs.2020.02.026
14. Pham BT, Bui KTT, Prakash I, Ly HB (2024) Hybrid artificial intelligence models based on adaptive neuro fuzzy inference system and metaheuristic optimization algorithms for prediction of daily rainfall. *Physics and Chemistry of the Earth, Parts A/B/C* 134: 103563. https://doi.org/10.1016/j.pce.2024.103563
15. Kaya E, Kaya A, Baştemur Kaya C (2025) Adaptive Network-Based Fuzzy Inference System Training Using Nine Different Metaheuristic Optimization Algorithms for Time-Series Analysis of Brent Oil Price and Detailed Performance Analysis. *Symmetry* 17: 786. https://doi.org/10.3390/sym17050786
16. Hao XH, Yang ZJ, Hao, Y (2021) Medium and long term prediction method of wind power based on improved gm-arma combination model. *Electrotechnical* 7: 11–13.
17. Fatemidokht H, Rafsanjani MK, Gupta BB (2021) Efficient and secure routing protocol based on artificial intelligence algorithms with UAV-assisted for vehicular ad hoc networks in intelligent transportation systems. *IEEE T Intell Transp Syst* 22: 4757–4769. https://doi.org/10.1109/TITS.2020.3041746
18. Yu T, Gan Q, Feng G, Han G (2022) A new fuzzy cognitive maps classifier based on capsule network. *Knowledge-Based Systems* 250: 108950. https://doi.org/10.1016/j.knosys.2022.108950
19. Sun L, Hou J, Xing CJ, Fang Z (2022) A robust hammerstein-wiener model identification method for highly nonlinear systems. *Processes* 10: 2664. https://doi.org/10.3390/pr10122664
20. Hou J, Su H, Yu C, Chen F, Li P (2022) Bias-correction errors-in-variables Hammerstein model identification. *IEEE T Ind Electron* 70: 7268–7279. https://doi.org/10.1109/TIE.2022.3199931

21. Meghraoui K, Racharak T, Ait El Kadi K, Bensiali S, Sebari I (2025) A New Integrated Neurosymbolic Approach for Crop-Yield Prediction Using Environmental Data and Satellite Imagery at Field Scale. *Artificial Intelligence in Geosciences* 6: 100125. https://doi.org/10.1016/j.aiig.2025.100125
22. Xu X, Lin Z, Li X, Shang C, Shen Q (2022) Multi-objective robust optimisation model for MDVRPLS in refined oil distribution. *Int J Prod Res* 60: 6772–6792. https://doi.org/10.1080/00207543.2021.1887534
23. Tang JN, Jin X, Zhang JZ, Lang YX, Liu JX (2017) Application analysis of dl/t 664-2016 code for application of infrared diagnosis of live equipment. *Smart grid* 5: 924–928.
24. Lauren J (2020) Support vector regression. In *Machine Learning*, 123–140. Academic Press: Cambridge, Massachusetts. https://doi.org/10.1016/B978-0-12-815739-8.00007-9
25. Abirami S, Pethuraj M, Uthayakumar M, Chitra P (2024) A systematic survey on big data and artificial intelligence algorithms for intelligent transportation system. *Case Stud Transp Pol* 17: 101247. https://doi.org/10.1016/j.cstp.2024.101247
26. Feng X, Xia WZ, Wang Z, Qian ZM, Liu J, Xu XQ (2021) Neural network water quality prediction of urban sewage plant based on SVR error compensation technology. *Water purification technology* 40: 92–98 + 158.
27. Selselejoo M, Ahmadifar HR (2025) DT-GWO: A hybrid decision tree and GWO-based algorithm for multi-objective task scheduling optimization in cloud computing. *Sustainable Computing: Informatics and Systems* 47: 101138. https://doi.org/10.1016/j.suscom.2025.101138
28. Wang C, Yang YL, Cai G, Zhang T (2024) Improvement of normalized prediction model of soil thermal conductivity. *Int Commun Heat Mass Transfer* 157: 107792. https://doi.org/10.1016/j.icheatmasstransfer.2024.107792
29. Kim YS, Kim MK, Fu N, Liu J, Wang J, Srebric J (2025) Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models. *Sustain Cities Soc* 118: 105570. https://doi.org/10.1016/j.scs.2024.105570
30. Fatemidokht H, Rafsanjani MK, Gupta BB, Hsu CH (2021) Efficient and secure routing protocol based on artificial intelligence algorithms with UAV-assisted for vehicular ad hoc networks in intelligent transportation systems. *IEEE T Intell Transp Syst* 22: 4757–4769. https://doi.org/10.1109/TITS.2020.3041746
31. Kumar N, Poonia V, Gupta BB, Goyal MK (2021) A novel framework for risk assessment and resilience of critical infrastructure towards climate change. *Technol Forecast Soc* 165: 120532. https://doi.org/10.1016/j.techfore.2020.120532
32. Sebastian A, Tantia V (2025) Multi-variate LSTM with attention mechanism for the Indian stock market. *International Journal of Information Management Data Insights* 5: 100350. https://doi.org/10.1016/j.jjimei.2025.100350
33. Zhao ML, Niu JR, Yang CG (2002) Discrimination of the concepts of temperature, relative temperature difference and temper-ature difference ratio in infrared detection technology. *Inner Mongolia electric power technology* 06: 8–10+15.

AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)