**Weigen CHEN[1], Xiaoping SU[1,2], Xi Chen[1], Qu Zhou[1], Hanguang Xiao[1]**

State Key Laboratory of Power Transmission Equipment and System Security and New Technology,Chongqing University,China(1)
Chengdu Electric Power Bureau,Chengdu,610041,China(2)

# Combination of Support Vector Regression with Particle Swarm Optimization for Hot-spot temperature prediction of oil-immersed power transformer

***Abstract.*** *The life expectancy and load capacity of oil-immersed power transformers are intimately associated with the winding hot spot temperature (HST). Thus, accurately predicting HSTs is essential in evaluating the life expectancy of power transformers and in preventing thermal failure. Previously, support vector machine (SVM) has been successfully employed to solve the regression problem of nonlinearity and small sample size. In the present study, given that the HSTs of transformers have a complex non-linear relationship with load information and environmental information, support vector regression (SVR) has been adopted to establish a model for the prediction of HSTs in power transformers. Among which, an improved particle swarm optimization (PSO) having passive congregation algorithm is utilized to determine the parameters of SVR. The PSO-SVR model has been applied to predict HSTs of a power transformer. Several experimental tests have been carried out involving a large power transformer in Sichuan Province, China, to verify the practicality and effectiveness of the proposed PSO-SVR model. In addition, PSO-SVR modeling results are compared with that of standard SVR and artificial neural network (ANN) by applying identical training and test samples. In conclusion, the PSO-SVR model has better prediction accuracy and generalization ability than both the standard SVR model and the ANN in the HST prediction of power transformers.*

***Streszczenie.*** *Artykul zajmuje się metodami przewidywania temperatury uzwojeń transformatora olejowego. W poprzedniej pracy do tego celu wykorzystano metodę SVM. W obecnej pracy uwzględniając nieliniowe zaleznośći od obciążenia i warunków zewnętrznych do przewidywania temperatury zastosowano metodę SVR (suport vector regression) oraz metodę optymalizacji wykorzystującą algorytmy mrówkowe.* ***(Wykorzystanie metody SVR oraz algorytmów mrówkowych do przewidywania temperatury uzwojeń transformatora olejowego)***

**Keywords:** power transformer, support vector regression, particle swarm optimization, parameters optimization, hot spot temperature, prediction,
**Słowa kluczowe:** transformator olejowy, temperatura uzwojen, metoda SVR

## Introduction

Oil-immersed power transformers are expensive and are difficult to replace after they have been integrated in the power system. The malfunctioning of these transformers results in serious losses to the safety, reliability, and operating costs of the power system. Power transformers suffer electric and thermal stresses from start to finish during operation. These stresses result in the deterioration of oil-immersed transformer insulating oil and other insulating materials, and may even cause the fire due to the damaged insulation. The maximum temperature which the winding and the insulation suffers, also known as winding hot spot temperature (HST), is the most direct factor that causes insulation deterioration and failure. Thus, monitoring and predicting HSTs facilitate the prevention of thermal incipient faults and the prediction of the life expectancy of power transformers [1, 2].

Presently, internal thermal characteristics analysis for power transformers is basically focused on the prediction of top oil temperature and HST [3, 4]. According to the top oil temperature rise model suggested by ANSI/IEEE C57.91 guide that is widely used in engineering, the steady-state top oil temperature of the transformer can be determined by calculating the top oil temperature rise over ambient temperature based on the first derivative model [5]. Examining dynamic models of transformer top oil temperature for an on-line monitoring and diagnostic system, the temperature data taken from large transformers in the field indicates that the IEEE model of top oil temperature rise over ambient temperature does not adequately account for daily variations in ambient temperature. Thus, a modified model that accurately predicts top oil temperature was later proposed [6]. For the sake of direct analysis, distributed measurements of transformer HST were carried out using thermocouples and fiber optic sensors. Nordman et al. addressed that HSTs of a large power transformer is not necessarily fixed to a particular phase, and claimed that overshoot exists [7]. Swift et al. derived a model for top oil temperature prediction using a thermal-electrical analogy method Based on the heat transfer theory, application of the lumped capacitance, and nonlinear thermal resistances [8]. An analogous thermal model and an equivalent circuit for HST determination were also presented. Qing et al. adopted neural networks to optimize the parameters of the differential equations to minimize the error of these analogy methods [9]. However, the accuracy of these calculations depends on whether the parameters are in line with the actual operation of the transformer. Additionally, the analogy model used the node and the lumped methods to study the average value of the characteristic temperature, rather than the true value. Although thermal-electrical analogy models ignore specific winding structure and thermal network approaches, these models take into account the disc arrangements, heat loss, and cooling ducts configuration and obtain satisfactory results in the prediction of HST. However, large power transformers are made from different materials and have different specifications, thermal structures, channel geometry, winding structures, and came from different manufacturers. Therefore, simulating HSTs for various power transformers using a common geometric model has its limitations. As a result, seeking a more effective and universal approach for the prediction of HST is essential.

The Support vector machine (SVM) which has been developed by Vapnik et al. in 1995, is a supervised learning approach based on statistical theory [10]. It has been gaining popularity due to its attractive features and promising empirical performance. An SVM based on the structural risk minimization (SRM) principle is able to control the complexity of the model and its generalization ability, which can be used for solving two-class or multi-class classification and regression problems in various fields, such as pattern recognition, function approximation, and time series prediction, among others [11]. SVM possesses many advantages including fast-learning, global optimization, and excellent generalization abilities due to minimizing the tradeoff between the complexity of the model and its generalization ability compared with other approaches, (i.e., artificial neural network, ANN). The modeling, including the SVM regression tasks, is usually referred to as support vector regression (SVR). Nowadays, SVM application in power systems is mainly concentrated in power transformer fault diagnosis and electric power load forecasting, among others [12, 13, 14]. In the current paper, given that the HST of the power transformer in service is a time series that actually changes with loadings, ambient temperature, and other varying parameters, the temperature data itself contains all the information about heat production and dissipation inside the transformer. Thus, HST prediction can be seen as a complex, non-linear function approximation problem between HST and its influencing factors. Hence, it is possible to predict transformer HST based on statistical theory using support vector regression (SVR). According to continuous online monitoring temperature characteristics, real-time load currents, ambient temperature, and other weather information from a large power transformer in Sichuan Province, a support vector regression (SVR) approach combined with particle swarm optimization (PSO) for its parameter optimization, is proposed to establish a model to predict the HST of power transformers. The validity and accuracy of the presented model has been verified with online measured data. In addition, the results of the PSO-SVR model were compared with those obtained using standard SVR and ANN.

**Brief Description on support vector regression(SVR)**

SVM combines the structural risk minimization (SRM) principle and the statistical machine learning theory (SLT). The modeling, including the SVM regression tasks, is usually referred to as support vector regression (SVR). The main purpose of SVR is to find a flat function $f(\mathbf{x})$ that can accurately predict the distribution of information. Suppose that there is a given training set of $N$ data points $(x_i,\, y_i)$ where $x_i$ denotes the input attribute and $y_i$ is the target value. SVR, the goal is to find a function $f(\mathbf{x})$ that has at most deviation from the actually obtained targets $y_i$ for all the training data and is, at the same time as flat as possible. The case of linear function $f$ is described as:

(1)
$$f(\mathbf{x}) = \mathbf{w}\cdot\mathbf{\Psi}(\mathbf{x}) + b \qquad i = 1, 2, \cdots, l$$
$$\mathbf{\Psi} : \mathbb{R}^n \to \mathbb{F}, \mathbf{w} \in \mathbb{F}$$

where, $\mathbf{\Psi}(\mathbf{x})$ denotes a nonlinear mapping function map $x_i$ from the input space $\mathbb{F}$ into a higher dimensional feature space $\mathbb{R}$, $\mathbf{w}$ is the weight vector and $b$ is the bias term. A small $\mathbf{w}$ means flatness in eq.( 1 ). Thus, minimizing the Euclidean norm is required, namely $\|\mathbf{w}\|^2$ . By introducing two slack variables $\xi$ and $\xi^*$ ,thus the coefficients $\mathbf{w}$ and $b$ can estimated by:

(2)
$$\begin{array}{ll} \text{min.} & \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \\ \text{st.} & \left\{ \begin{array}{l} y_i - [\mathbf{w}\cdot\mathbf{\Psi}(x_i)] - b \le \varepsilon + \xi_i \\ b - y_i + [\mathbf{w}\cdot\mathbf{\Psi}(x_i)] \le \varepsilon + \xi_i^* \\ \xi_i \ge 0, \quad \xi_i^* \ge 0 \end{array} \right. \end{array}$$

where, $C$ is a regularized factor, the constant $C > 0$ determines the tradeoff between the flatness of $f$ and the amount up to which deviations larger than $\varepsilon$ are tolerated.This condition corresponds to dealing with a so called $\varepsilon$-intensive loss function $L_\varepsilon(y_i, f(x_i))$, which is equal to zero if the error of the forecasting value is less than $\varepsilon$, otherwise the loss is equal to the value beyond $\varepsilon$. Furthermore,introducing a Lagrange function:

(3)
$$\begin{aligned} L(\mathbf{w}, \xi_i, \xi_i^*) \; &= \tfrac{1}{2}\|\mathbf{w}\|^2 + C\textstyle\sum_{i=1}^{l}(\xi_i - \xi_i^*) \\ &- \textstyle\sum_{i=1}^{l}(\eta_i\xi_i + \eta_i^*\xi_i^*) \\ &- \textstyle\sum_{i=1}^{l}\alpha_i(\varepsilon + \xi_i - y_i + \mathbf{w}\cdot\mathbf{\Psi}(\mathbf{x}) + b) \\ &- \textstyle\sum_{i=1}^{l}\alpha_i^*(\varepsilon + \xi_i^* + y_i - \mathbf{w}\cdot\mathbf{\Psi}(\mathbf{x}) - b) \end{aligned}$$

where$\eta_i$, $\eta_i^*$, $\alpha_i$ , and $\alpha_i^*$ are the Lagrange multipliers. Following the saddle point condition, the partial derivatives of $L$ with respect to the primal variables ($\mathbf{w}$, $b$, $\xi_i$, $\xi_i^*$ ) have to vanish for optimality. Then the optimization problem can be rewritten as:

(4)
$$\begin{array}{ll} \text{max.} & \Big\{ -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(\mathbf{\Psi}(\mathbf{x}_i), \mathbf{\Psi}(\mathbf{x}_j) \\ & \quad -\varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) \Big\} \\ \text{st.} & \sum_{i=1}^{l}(\alpha_i + \alpha_i^*) = 0 \quad \alpha_i, \alpha_i^* \in [0, C] \end{array}$$

An optimal desired weight vector of the regression hyper plane is obtained using the Lagrange multipliers and those calculated. $\mathbf{w}$ can then be rewritten as:

(5)
$$\mathbf{w} = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\cdot\phi(\mathbf{x}_i)$$

Hence, the regression function has the following explicit form:

(6)
$$f(\mathbf{x}) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(\mathbf{x}_i, \mathbf{x}) + b$$

where $K(\mathbf{x}_i, \mathbf{x}) = \phi(\mathbf{x}_i)\cdot\phi(\mathbf{x})$, is a kernel function which value equals the inner product of $\phi(\mathbf{x}_i)$ and $\phi(\mathbf{x})$ . Choosing different types of kernel functions can generate different SVR models. The linear kernel, radial basis function (RBF), polynomial basis function, and sigmoid function are commonly used. Noticed that only one variable needs to be determined, the radial basis function,$K(\mathbf{x}_i, \mathbf{x}) = \exp(-\gamma\|\mathbf{x}_i - \mathbf{x}\|^2)$, was used in the current work, and then the user-determined parameters of the regression problem are $\varepsilon$, $C$ and $\gamma$.

**Parameters selection of SVM based on PSO**

As stated above, the three free parameters, i.e., $\varepsilon$, $C$ and $\gamma$, greatly affect the generalization ability of SVR. However, no prior information is available about the appropriate values of the parameters. Therefore, a key step is to search the optimal parameter subset ($\varepsilon$, $C$, $\gamma$) for SVR. The standard particle swarm optimizer (PSO) is a population-based algorithm which was inspired by the social behavior of animals, such as fish schooling and bird flocking. PSO can not only solve a variety of difficult optimization problems, but also have shown a faster convergence rate than other evolutional algorithms(EA). Although PSO may outperform other EAs in the early iterations, its performance may not be as competitive when the number of generations is increased [2]. In the present paper, an improved particle swarm optimizer with passive congregation is employed to search global solutions of the optimal parameters ($\varepsilon$, $C$, $\gamma$) using on-site measurements for SVR.

PSO performs searches using a population (called swarm) of individuals (called particles) that are updated from iteration to iteration, during the iterations, it searches the best parameter by regulating velocity and location of particles. Each of particle swarm is made up of a parameter vector ($\varepsilon$, $C$, $\gamma$). For the particle at iteration k, it has the following two attributes: one is current position in the 3-dimensional search space $\mathbf{u}_i^k = (\mathbf{u}_{i1}^k, \mathbf{u}_{i2}^k, \mathbf{u}_{i3}^k)$, the other attribute is the corresponding velocity $\mathbf{v}_i^k = (\mathbf{v}_{i1}^k, \mathbf{v}_{i2}^k, \mathbf{v}_{i3}^k)$ . Noticed that a population may lose diversity and is more likely to confine the search around local minima if committed too early in the global best found so far. In order to overcome this factor, an improved PSO with passive congregation has been developed. In each iterative process, the swarm can be updated by the following equations:

$$\text{(7)}\qquad \begin{aligned} \mathbf{v}_i^{k+1} &= \omega \mathbf{v}_i^k + c_1 r_1 \cdot (\mathbf{p}_{ibest}^k - \mathbf{u}_i^k) \\ &+ c_2 r_2 \cdot (\mathbf{p}_{gbest}^k - \mathbf{u}_i^k) \\ &+ c_3 r_3 \cdot (\mathbf{R}_i^k - \mathbf{u}_i^k) \\ \mathbf{u}_i^{k+1} &= \mathbf{u}_i^k + \beta \cdot \mathbf{v}_i^{k+1} \end{aligned}$$

where: $\mathbf{R}_i^k$ is a particle randomly selected from the swarm, $\mathbf{p}_{ibest}^k$ is the local best position of the $i^{th}$ particle and $\mathbf{p}_{gbest}^k$ denotes the global best position among all the particles in the swarm; $k$ is the evolutional generation, $\omega$ , the inertial weight, is used to balance the global exploration and local exploitation and consequently results in a better optimum solution. $r_1$, $r_2$ and $r_3$ are elements from two uniform random sequence in the range [0, 1]; Positive constant $c_1$ and $c_2$ are both acceleration constants, which also control how far a particle will move in a single iteration, whereas $c_3$ denotes the passive congregation coefficient and $\beta$ is constraint factor used to control the velocity weight, whose value is usually set to 1.The process of optimizing the SVM parameters with PSO can be described more detail below:

Step 1: Initialization. Set $k$ =0, randomly initialize positions and corresponding velocities;

Step 2: Checking feasibility. Check the feasibility of the $i^{th}$ particle in the swarm. Then reset $\mathbf{u}_i$ to the previous feasible position $\mathbf{u}_i^{k-1}$ if the current position $\mathbf{u}_i^k$ beyond the feasible region;

step3: Evaluating fitness value. In this study, the $k$-fold cross validation is used to evaluate fitness. and root mean square error (RMSE), which directly reflects the regression performance of SVR, is selected as fitness function:

Step 4: Updating the global and personal best according to the fitness evaluation results. comparing the fitness value of $p_{ibest}^k$ with the corresponding objective function, set $\mathbf{p}_{ibest}^k$ to the current position if the corresponding objective function better than the fitness value of $\mathbf{p}_{ibest}^k$ ; similarly, if the objective function is also better than the fitness value of $\mathbf{p}_{gbest}^k$, then set $\mathbf{p}_{gbest}^k$ to the current position $\mathbf{u}_i^k$ .

Step 5: Updating $\mathbf{R}_i$. Randomly select a particle from the swarm as $\mathbf{R}_i$ .

Step 6: Calculating velocities. The particle flies toward a new position by calculating the velocity of position change. The velocity of each particle is calculated by eq.(7).

Step 7: Updating position value. Each particle moves to its next position according to eq.(7).

Step 8: Termination. Repeating the same procedures from Step 2 to Step 7 until termination conditions met.

**SVR modeling for hot spot temperature prediction**

To verify the validity and to show the prediction accuracy of the presented PSO-SVR model, the results of the PSO-SVR model are compared with that of standard SVR, ANN. Based on the existing various features information of online monitoring and heat run test of power transformer in service, temperature characteristics, real-time load currents, ambient temperature, and other weather information of a large power transformer (750MVA/500kV) in Sichuan Province from September 25, 2010 to October 03, 2010 were collected as the input and output variables of the PSO-SVR model. Each characteristic features was recorded every hour during data acquisition, thus the dataset contains a total of 192 samples. Each sample contains top oil temperature(TOT), power losses($P_L$), reactive losses($R_L$), load current($I$), ambient temperature($T_{amb}$), humidity($H$) and wind speed($v_w$).

Table 1. Partial input/output data samples.

| No. | $P_L$ (MW) | $R_L$ (Mvar) | $I$ (A) | $T_{amb}$ (°C) | $H$ (%) | $v_w$ (km/h) | TOT (°C) |
|---|---|---|---|---|---|---|---|
| 1 | 166.54 | -12.86 | 431.72 | 18 | 83 | 14.4 | 44.2 |
| 2 | 190.34 | -6.43 | 509.96 | 17 | 77 | 18 | 43.6 |
| 3 | 174.26 | -22.51 | 433.12 | 16 | 88 | 14.4 | 43.8 |
| 4 | 162.04 | -11.57 | 407.81 | 15 | 94 | 18 | 44.6 |
| 5 | 160.76 | -5.14 | 416.25 | 17 | 82 | 10.8 | 44.8 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 188 | 155.61 | -16.72 | 388.12 | 17 | 88 | 14.4 | 45.9 |
| 189 | 160.11 | -18.65 | 402.19 | 17 | 88 | 18 | 46.9 |
| 190 | 178.76 | -14.79 | 448.59 | 19 | 78 | 18 | 44.8 |
| 191 | 193.55 | -10.29 | 502.23 | 20 | 73 | 7.2 | 43.5 |
| 192 | 173.62 | -25.72 | 433.12 | 19 | 73 | 10.8 | 43.8 |

A total of 7 features were employed as the input variable. Meanwhile, the expected HST were recorded every hour corresponding to each sample and acted as the output variable. Supposed the input vector as **X** ($7 \times 192$ dimensions) and the output vector as **Y** ($1 \times 192$), respectively. In the current study about HST predicting, three cases were chosen to verify the performance of the model: Case 1 (5:3), the dataset was divided into eight data partitions, and then the first five partitions were gathered to constitute an initial training data fraction (**X**$1$, **Y**$1$), whereas the remaining three partitions were used as an initial testing data fraction (**X**$2$, **Y**$2$); Case 2 (6:2), six out of the eight partitions were chosen as the training data (**X**$3$, **Y**$3$) and the two remaining partitions acted as testing data (**X**$4$, **Y**$4$). In Case 1, the prediction model was determined using the training dataset **X**$1$ first, then PSO algorithm was introduced to obtain the optimal parameters ($\varepsilon$, $C$, $\gamma$) by minimizing the fitness function of PSO. The optimal parameters are then utilized to train the SVM model. Finally, testing data sets **X**$2$ are used to examine the accuracy of the HSTs prediction. In the modeling process, 5-fold and 6-fold cross validation was used to evaluate the fitness for Case 1 and Case 2, respectively. In addition, normalization of the data before training is essential to eliminate all the interference of different dimensions of various inputs on SVR, which can improve the generalization capability of PSO-SVM, ANN. The normalized data of HSTs are shown in Fig. 1. Different training data sets and testing data sets are constructed in each case. The partial input/output data samples are given in Tab. 1. Moreover, during the SVR parameters identification by PSO, the primary principle is that of choosing the ones with the highest accuracy rate. If various solutions correspond to the highest accuracy rate, then the ones that correspond to a relatively small regularized factor $C$ is selected; and if there are various parameters corresponding to the minimum regularized factor, then the first group searched parameters corresponding to the minimum regularized factor is selected as the optimal SVR parameters. The four following functions have been used to measure the model generalization performance based on extensive tests for PSO-SVR and a number of modeling experiences: the maximum absolute percentage error (MPE), mean absolute error (MAE), mean absolute percentage error (MAPE), and correlation coefficient ($R^2$ ), which is defined by : $R^2 = \sum_{j=1}^{n} (\hat{y}_j - \bar{y})^2 / \sum_{j=1}^{n} (y_j - \bar{y})^2$.

**Experimental analysis and discussion**

In PSO-SVR, the training data set are fed into the PSO-SVR model, and the structural risk minimization principle is employed to minimize the training error. The three kernel

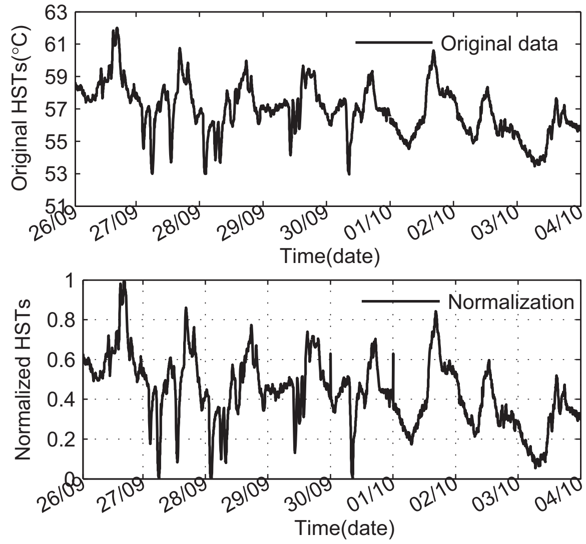


Fig. 1. Original and Normalized data of HSTs

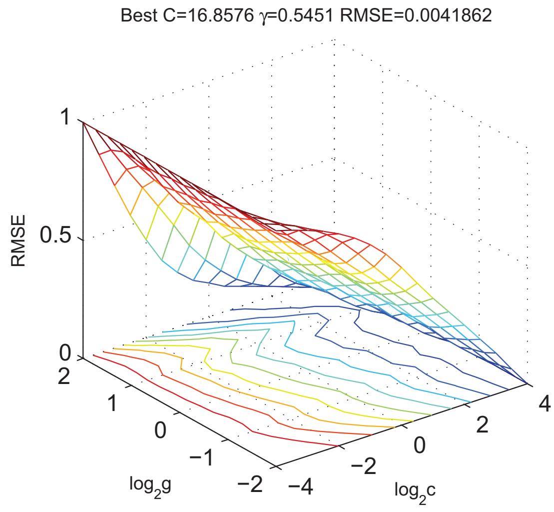


Fig. 2. Parameters optimization results by PSO (Case 1)

parameters ($\varepsilon$, $C$, $\gamma$) of the PSO-SVR model adjusted by P-SO are employed to calculate the validation error while the improvement in training errors occurs. The adjusted parameters with the minimum validation errors are then selected as the most appropriate parameters. The most suitable kernel parameters in the PSO- SVR SVM models are those with the smallest testing MAPE value. In training PSO-SVR for predicting HST, the population was set as 20, the limit factor of displacement is taken as $c1 = c2 = 2$, inertia weight, which is typically chosen in the range of [0,1], is set here down from 0.9 to 0.4 linearly, and the termination algebra is 200. Moreover, root mean square error (RMSE) was selected as the fitness function of PSO. The parameters optimization result of Case 1 is shown in Fig. 2. In the current study, a recursive mapping ANN of 7×9×1 recursive form was also configured to describe transformer thermal behavior for predicting HST. For the sake of fairly comparing the performance of the proposed model with that of standard SVR, ANN, the learning and test processes were carried out directly in terms of the same training and test samples. Comparison of the forecasting results among PSO-SVR, standard SVR, and ANN are listed Tab. 2. The standard SVR and ANN are listed from the RMSE, MAE, MAPE, and $R^2$ of the prediction results for two different cases estimated by PSO-SVR, as shown in Tab. 2. Obviously, the regression performance of PSO-SVR models surpasses that of both the standard SVRs and ANNs with their smallest RMSE/MAE/MAPE indexes among these three models under the same training and test conditions. At the same time, Tab. 2 also lists the correlation coefficients ($\varepsilon$, $C$, $\gamma$) for the three different cases. The correlation coefficients of PSO-SVR are all larger than the other models. Hence, the regression effect fitted by the proposed PSO-SVR model is better than that of standard SVR and ANN. Fig. 3 and Fig. 4 describe the results of Case 1, which were obtained from the PSO-SVR, standard SVR, and ANN, respectively. The ANN model only represents the input and output relationship accurately within the range covered by the training data when the recursive ANN mapping is used, as shown in Fig. 3 and Fig. 4. Outside this range, the model response and measurements do not agree with each other satisfactorily, especially under rapid variations of load. Notably, the prediction results using the standard SVR with randomly slected parameters are still closer to the measuments although the maximum training error of the standard SVR model is larger than that of ANNs. In contrast, the proposed PSO-SVR model not only

Table 2. Comparison of the predicting results of two different cases among PSO-SVR, standard SVR and ANN.

| | Case 1 | | | Case 2 | | |
|---|---|---|---|---|---|---|
| | SVR | ANN | PSOSVR | SVR | ANN | PSOSVR |
| MAE(°C) | 0.854 | 1.248 | 0.670 | 0.591 | 1.237 | 0.583 |
| MAPE(%) | 1.436 | 2.315 | 1.045 | 1.014 | 2.228 | 0.977 |
| RMSE(°C) | 0.920 | 1.319 | 1.186 | 1.142 | 1.307 | 0.902 |
| $R^2$ | 0.891 | 0.817 | 0.976 | 0.905 | 0.842 | 0.983 |

exhibited a comparable performance with ANNs in the training process, but also had the smallest predicting error for the testing data among the three models using an improved PSO algorithm with passive congregation to optimize the SVR parameters.

Fig. 5 shows the training set $\mathbf{X}1$ (7×120 dimensions) and the test set $\mathbf{X}2$ (7×72 dimensions) being placed into the SVR model by the PSO optimization parameters and corresponding output result of the HST model. The MAE, RMSE, and MAPE of the PSO-SVR model are 0.67°C, 1.186 °C, and 1.045%, respectively, compared with that of the corresponding online monitoring temperature. Additionally, Fig. 5 illustrates the maps of PSO-SVR model prediction results corresponding to the substation on-site monitoring data to focus on analyzing the real-time HST prediction capabilities of the PSO-SVR mode. Evidently, the data points located are within the vicinity of 45° line evenly with the correlation coefficient $R^2$=0.976024. This condition intuitively shows the good agreement between the predicted and the measured values. The satisfactory agreements between the model outputs and the real-time measurements for Cases 2 are also illustrated in Tab. 2. This condition also demonstrates that increasing the number of the training data samples facilitates the regression accuracy, thus indicating that PSO-SVR can be successfully utilized in timely HST prediction. In addition, the PSO-SVR model still obtained an accurate representation of the actual thermal states with the untrained datasets, whereas the data samples in all three cases contain both the trained and the untrained dataset. Therefore, the proposed PSO-SVR model not only exhibited good learning abilities, but also presented a strong generalization performance in predicting HST. Hence, the PSO-SVR model provides a new method for real-time HST prediction in oil-immersed power transformers.

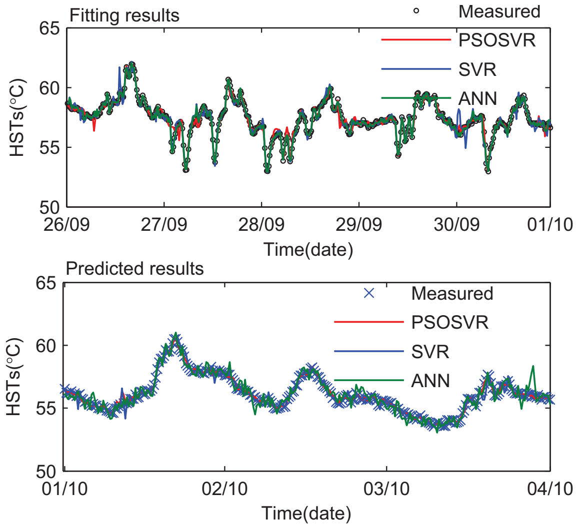


Fig. 3. HST predicting values of PSO-SVR, SVR, and ANN (Case1)

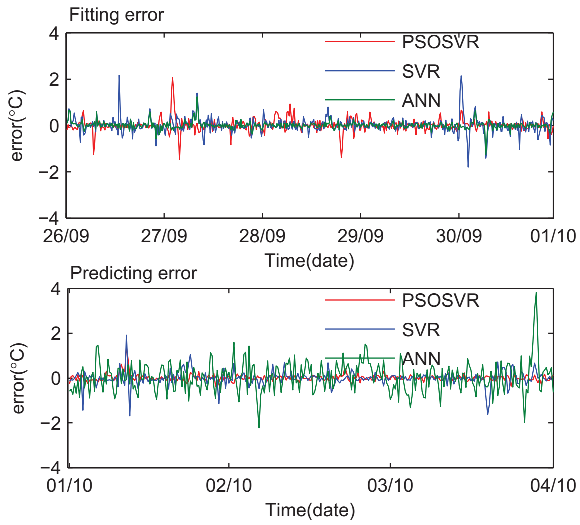


Fig. 4. Predicting error of PSO-SVR, SVR, and ANN (Case1)

## Conclusions

In the current study, given that HSTs of transformers have a complex non-linear relationship with load information and environmental information, support vector regression (SVR) has been adopted to establish an HST prediction model for power transformers. Among which, an improved particle swarm optimization (PSO) with passive congregation algorithm was used to identify the appropriate parameters of SVR. Real datasets sampled from a large power transformer are used to validate the effectiveness and feasibility of the proposed PSO-SVR model in HST prediction. Moreover, the PSO-SVR modeling results were compared with that of the standard SVR and ANN modesl. The prediction performance of PSO-SVR surpasses that of both standard SVR and ANN with their smallest RMSE/MAE/MAPE indexes under the same training and test conditions. The prediction accuracy of PSO-SVR model can also be further improved by increasing the number of training samples. In conclusion, the proposed PSO-SVR model is applicable and practical in accurately predicting HSTs of power transformers.

*Project supported by the Funds for Innovative Research Groups of China(51021005).*

BIBLIOGRAPHY

[1] Amoiralis E., Tsili M., Kladas AG.: Transformer Design and Optimization: A Literature Survey, *IEEE Transactions on Power Delivery*, 24(4), pp. 1999-2024, 2009.

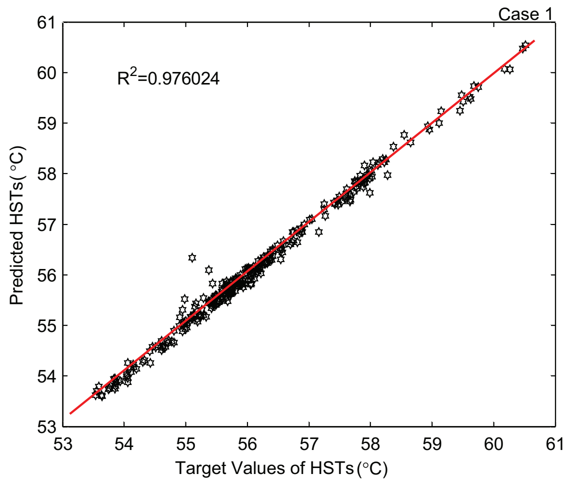


Fig. 5. Pair wise comparison of PSO-SVR results vs. measured HSTs (Case1)

[2] Tang W.H., Wu Q.H.: Condition Monitering and Assessment of Power Transformers using Computational Intelligence, *NewYork: Springer*, 1st Edition, 2011.
[3] Weigen C., Chong P., Yuxin Y.: Power transformer top-oil temperature model based on thermal–electric analogy theory, *European Transactions On Electrical Power*, 19(3), pp. 341-354, 2009.
[4] Han Y., Song Y. H.: Condition monitoring techniques for electrical equipment-a literature survey, *IEEE Transactions on Power Delivery*, 18(1), pp. 4-132, 2003.
[5] IEEE, Std 1538-2000.: IEEE Guide for determination of maximum winding temperature rise in liquid-filled transformers, *IEEE* , 2000.
[6] Lesieutre B., Hagman W.: An improved transformer top oil temperature model for use in an on-line monitoring and diagnostic system, *IEEE Transactions on Power Delivery*, 12(1), pp. 249-256, 1997.
[7] Nordman H., Rafsback N.T., Susa D.: Temperature responses to step changes in the load current of power transformers, *IEEE Transactions on Power Delivery*, 18(4), pp. 1110 - 1117, 2003.
[8] Swift G., Molinski T.S., Lehn W.: A fundamental approach to transformer thermal modeling. I. Theory and equivalent circuit, *IEEE Transactions on Power Delivery*, 16(2), pp. 171-175, 2001.
[9] He Q., Si J., Tylavsky D.J.: Prediction of top-oil temperature for transformers using neural networks, *IEEE Transactions on Power Delivery*, 15(4), pp. 1205-1211, 2000.
[10] Vapnik V.N.: The nature of statistical learning theory, *NewYork: Springer*, 2nd Edition, 1999.
[11] Wen Y.F., Cai C.Z., Liu X.H.,Pei J.F., Zhu X.J., Xiao T.T.: Corrosion rate prediction of 3C steel under different seawater environment by using support vector regression, *Corrosion Science*, 51(2), pp. 349-355, 2009.
[12] Fei S.W., Wang M.J., Miao Y.B.: Particle swarm optimization-based support vector machine for forecasting dissolved gases content in power transformer oil, *Energy Conversion and Management*, 50(6), pp. 1604-1609, 2009.
[13] Ganyun L.V., Cheng H., Zhai H., Dong L.: Fault diagnosis of power transformer based on multi-layer SVM classifier, *Electric Power Systems Research*, 74(1), pp. 1-7, 2005.
[14] Yunqiu W.,Zongxi T., Yuehang X., Yunchuan G., Biao Z.: SVR for Measuring Electromagnetic Parameters, *IEEE Transactions on Magnetics*, 43(12), pp. 4071-4075, 2007.

***Authors:*** Prof. Weigen Chen, State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing, 400044,China, email: weigench@gmail.com, Ph.D. Xiaoping Su, State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing, 400044, China, Chengdu Electric Power Bureau, Chengdu, 610041, China, email: cqusuxiaoping@gmail.com