Transformer Top-oil Temperature Modeling Based on Kernel-based Extreme Learning Machine

Hua HUANG¹, Ben-gang WEI¹,*, Xiao-wu Qi², Yan-shun Xu², Shuang Hu², Kai-qi Sun², Mei-yan Wang² and Jing Guo²

¹Equipment State Evaluation Center, Shanghai Municipal Electric Power Company EPRI, Shanghai, 200437, China
²School of Electrical Engineering, Shandong University, Jinan 250061, China

*Corresponding author

Keywords: Power transformer, Top-oil temperature, Kernel-based extreme learning machine, Particle swarm optimization.

Abstract. Transformer top-oil temperature (TOT) and winding hot-spot temperature (HST) are key indices to evaluate thermal condition of transformers. In order to improve TOT prediction accuracy, a TOT prediction model based on kernel-based extreme learning machine is established and particle swarm optimization algorithm is adopted to train the model and optimize the kernel parameters. The proposed model is tested on a 50MVA 110/37kV ONAN transformer. Besides, to verify the advantages of the proposed model, it’s compared with several traditional data-driven models. The results demonstrate the validity and accuracy of the proposed model.

Introduction

The power transformer dynamic loading capacity and insulation aging speed are mainly decided by its internal thermal condition, of which the HST and TOT are the typical indicators [1]. Therefore, monitoring the HST and TOT is important to ensure the thermal safety of the transformers [1-2]. At present, direct measurement of HST has not been widely used due to difficult installation and maintenance of winding temperature optical fiber sensors [3]. Generally, HST is evaluated via TOT, which is measured for most transformers [1-2]. Hence, accurate modeling and predicting TOT in advance is a practical solution for transformer thermal condition estimation and it helps to fully explore loading capacity of the units and keep the units operate safely.

There have been several existing methods to predict transformer TOT and HST. According to their principles, these methods are divided into two categories: semi-physical models and data-driven models. The representatives of semi-physical models are the methods recommended by IEEE and IEC loading guide [1-2], and the thermal circuit models [4-6]. The parameters of semi-physical models are based on transformer temperature rise test report and such models are less in accuracy than models trained from measured data [7]. More recently, nonlinear fitting regression techniques have been used to produce more accurate TOT models, such as artificial neural networks (ANN) models, least square support vector machine (LS-SVM) models, and so on [8-9].

The authors propose a data-driven model based on kernel-based extreme learning machine (KELM). As for most existing data-driven models, the prediction result may be smaller than the real one, which may lead to over-optimistic estimation of TOT. In this paper, particle swarm optimization algorithm (PSO) is adopted to train the KELM and optimize the kernel parameters. In order to achieve conservative estimation of TOT, during the training process, the positive error and negative error is processed with different tolerance.

The proposed model is trained and tested and the results demonstrate the validity and accuracy of the proposed model. Besides, the proposed model is compared with BP ANN model and LS-SVM model, the results show that the proposed model can achieve conservative estimation of TOT and the prediction accuracy is better.
Theory Basis

Kernel-based Extreme Learning Machine

Kernel-based extreme learning machine (KELM) is a kind of kernel-based feed forward neural network (SLFN), which works well in nonlinear regression and classification. The output of KELM is given by (1) [10]:

\[ f(x) = [K(x, x_1), ..., K(x, x_n)] \beta \]

where \( K(x, x_i) \) is the kernel function and Gaussian kernel function is chosen:

\[ K(x, x_i) = \exp\left(-\frac{\|x-x_i\|^2}{\sigma^2}\right) \]  

(2)

where \( \sigma \) is kernel parameter. \( \sigma \) is optimized by PSO during the KELM training process. Fig. 1 displays the symbolic KELM model.

![Figure 1. Structure of KELM model.](image)

To achieve conservative prediction result and optimize the kernel parameters, the model is trained by PSO.

PSO algorithm

The PSO algorithm is a swarm intelligence optimization algorithm, which works well in engineer optimization [11, 12]. The position of each particle represents a feasible solution \( u \), and the velocity represents particle shift step length in each iteration. The optimization procedure is implemented by particle velocity and position updates. Firstly, a certain number of particles are generated randomly. Then, the velocity and position of each particle are updated according to (3) and (4) [12]:

\[ v_i(t+1) = w \cdot v_i(t) + c_1 \cdot rand() [P_{best,i} - p_i(t)] + c_2 \cdot rand() [G_{best,i} - p_i(t)] \]

(3)

\[ p_i(t+1) = p_i(t) + v_i(t+1) \]

(4)

where \( v_i(t) \) and \( p_i(t) \) are the \( i \)th dimensional velocity and position in iteration \( t \); \( w \) is inertia weight; \( c_1 \) and \( c_2 \) are acceleration coefficients; \( rand() \) is random numbers between 0 and 1; \( P_{best,i} \) and \( G_{best,i} \) are individual optimal solution and global optimal solution. When the updating steps reach the maximum number, the final global optimal solution is set to be the best solution.

Transformer Top-oil Temperature Modeling

Inputs of the KELM Model

As aforementioned, the proposed KELM TOT model is data-driven, which means that the relation between TOT and its influence factors is treated as nonlinear mapping [7-9]. Therefore, the main influence factors, the load current (\( I \)) and ambient temperature (\( \theta_{amb} \)) are chosen as the input of the model. As the TOT change lags behind the change of influence factors such as load current, the temporal delay lines which refer to samples of previous moment must be considered. Based on existing references, the number of delay lines is chosen to be 2 (the sample interval is 15min) [8-9]. Therefore, the inputs of model are \( I(n), I(n-1), I(n-2), \theta_{amb}(n), \theta_{amb}(n-1), \theta_{amb}(n-2) \).
The output TOT $\theta_{oil}$ is given by:

$$\theta_{oil}(x) = \left[ K(x,x_1), \ldots, K(x,x_N) \right] \cdot \beta$$

(5)

where $x$ is the input vector constituted by the input variables, $x_1, x_2, \ldots x_N$ are input training data.

**Model Training and Kernel Parameters Optimization**

Most existing data-driven models do not consider the difference of positive and negative prediction error. The negative prediction error may lead to over-optimistic estimation of TOT, which is unfavorable for transformer thermal condition estimation. On the contrary, positive prediction error (the prediction result $\theta_{oil,e}$ is larger than the real one $\theta_{oil}$) leads to conservative transformer thermal condition estimation, which is helpful for transformer safe operation. Therefore, the positive error should be processed with larger tolerance. In this paper, the KELM training and kernel parameters optimization is implemented simultaneously via PSO. The particle is constituted by KELM output weights vector $\beta$ and kernel parameters $\sigma$:

$$u = [\beta, \sigma]$$

(6)

The mean square error (MSE), given by (7), is the objective function:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (\theta_{oil} - \theta_{oil,e})^2.$$  

(7)

The problem is reduced to be a single-objective problem: finding optimal $u^*$ to minimize $MSE (u)$, and the constraint is that the prediction result must be larger than the measured TOT ($\theta_{oil,e} \geq \theta_{oil}$).

During the updating process, if a particle violates the constraint, the particle is discarded, and a new particle is generated randomly. The procedure of model training and kernel parameter optimization is shown by Fig.2.

1. The measured datasets $D=\{x_n, y_n\}$ are divided into training set $D_{train}$ and test set $D_{test}$.
2. The criteria to update the solution is $MSE$, and the particles which do not satisfy the hard constraint are discarded, and new particles are generated as replacements.
3. When the updating steps reach the maximum number, the final global optimal solution is set to be the best solution.

**Case Study**

The measured data of a 50MVA 110/37kV ONAN transformer acquired in a substation are adopted to conduct the simulation test, the load current ($I$), ambient temperature ($\theta_{amb}$), and top oil ($\theta_{oil}$) pattern are relative to six-month observation period (January to March and October to December).
with 15-minute sampling interval. The example data of 5-day observation period in January is shown in Fig.3. The data in in December are $D_{train}$, and the rest are $D_{test}$.

**Model Training and Kernel Parameters Optimization**

The PSO is adopted to train the KELM and optimize the kernel parameters. The inertia weight $w=0.6$, the acceleration coefficient $c_1=2$, $c_2=2$. The number of particles is 40, and the maximum iteration number is 200. Fig.4 shows the fitness variation in relation to iterations. It can be observed that after about 500 iterations, the model training and parameters optimization process could be considered to be terminated. The optimized $\sigma$ is 24.173.

![Figure 3. Measure data example in January.](image)

![Figure 4. Convergence curve of PSO.](image)

**Prediction Results and Analysis**

The trained models are used to predict TOT by (5). The 5-month $D_{test}$ is used to verify the validity of the model. Fig.5 displays the comparison of the prediction results and the measured TOT of 5-day period. The mean square error ($MSE$) is 0.51 and the maximum error ($\epsilon_{\max}$) is 1.66.

![Figure 5. TOT prediction results and measured TOT.](image)

As it can be noted from Fig.5, the TOT prediction results fit well with the measured TOT, and the prediction results are slightly larger than the measured TOT, which means that the prediction results are conservative estimation of the TOT.

**Comparison with BP Network and LS-SVM**

In order to demonstrate the model performance, the proposed PSO-KELM model is compared with BP neural network and LS-SVM. Through lots of trials and comparisons, the best settings of BP network are that the node numbers of input layers, hidden layers, and output layers are separately 6, 14 and 1. The hidden layer exciting function of BP network is sigmoid function and the output layer transfer function is logsig function. The learning algorithm is chosen to be Levenberg-Marquardt algorithm. The kernel function of LS-SVM is the same with KELM and the kernel parameters are
also optimized via PSO algorithm. The prediction results of the combination models using PSO-KELM, BP neural network and LS-SVM are shown in Tab.1.

<table>
<thead>
<tr>
<th></th>
<th>PSO-KELM</th>
<th>BP Network</th>
<th>LS-SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.51</td>
<td>1.19</td>
<td>0.75</td>
</tr>
<tr>
<td>$e_{\text{max}}$(ºC)</td>
<td>1.66</td>
<td>2.63</td>
<td>2.34</td>
</tr>
</tbody>
</table>

As is shown in Tab.1, the $MSE$ of PSO-KELM is separately 42.86%, 68% of the BP network and LS-SVM, and the $e_{\text{max}}$ of PSO-KELM is separately 63.12% and 70.94% of the BP network and LS-SVM. Hence, the PSO-KELM can get higher accuracy in predicting transformer TOT.

Fig.6 displays the prediction error of the proposed PSO-KELM model and LS-SVM model with a 5-day period. It is obvious that the prediction error of PSO-KELM model is positive while the sign of LS-SVM prediction error is uncertain. This means that the PSO-KELM model could achieve conservative estimation of TOT while the LS-SVM may lead to over-optimistic estimation of TOT. Therefore, the PSO-KELM model suits transformer thermal estimation better.

![Figure 6. TOT prediction error of PSO-KELM and LS-SVM.]

**Conclusion**

A transformer TOT prediction model, based on kernel-based extreme learning machine and PSO algorithm, is proposed in this paper. KELM model is used to map the TOT and its influence factors, and PSO is used for KELM training and kernel parameters optimization. During training procedures, in order to achieve conservative estimation of TOT, the positive error is processed with larger tolerance. Compared with BP neural network and LS-SVM, the prediction error of KELM is smaller and the prediction results are conservative. The proposed TOT prediction model in this paper improves the accuracy in transformer thermal calculation and prediction, which is conducive to guiding transformer loading and the deep mining of transformer monitoring information.

**References**


