



Article

A New Prediction Model for Transformer Winding Hotspot Temperature Fluctuation Based on Fuzzy Information Granulation and an Optimized Wavelet Neural Network

Li Zhang, Wenfang Zhang *, Jinxin Liu , Tong Zhao , Liang Zou and Xinghua Wang

School of Electrical Engineering, Shandong University, Jinan 250061, China; zhlee@sdu.edu.cn (L.Z.); ljxwillingheart@163.com (J.L.); zhaotong@sdu.edu.cn (T.Z.); zouliang@sdu.edu.cn (L.Z.); Wangxinghua@sdu.edu.cn (X.W.)

* Correspondence: zwf02023@163.com; Tel.: +86-0531-81696129

Received: 27 October 2017; Accepted: 21 November 2017; Published: 1 December 2017

Abstract: Winding hotspot temperature is the key factor affecting the load capacity and service life of transformers. For the early detection of transformer winding hotspot temperature anomalies, a new prediction model for the hotspot temperature fluctuation range based on fuzzy information granulation (FIG) and the chaotic particle swarm optimized wavelet neural network (CPSO-WNN) is proposed in this paper. The raw data are firstly processed by FIG to extract useful information from each time window. The extracted information is then used to construct a wavelet neural network (WNN) prediction model. Furthermore, the structural parameters of WNN are optimized by chaotic particle swarm optimization (CPSO) before it is used to predict the fluctuation range of the hotspot temperature. By analyzing the experimental data with four different prediction models, we find that the proposed method is more effective and is of guiding significance for the operation and maintenance of transformers.

Keywords: transformer winding; hotspot temperature; prediction model; fuzzy information granulation; wavelet neural network

1. Introduction

The power transformer has always been the focus of monitoring and protection as an important piece of power transmission equipment. The main reason for the ending of its service life is the loss of proper insulating capacity. The transformer winding hotspot temperature (TWHT), which is defined as the temperature of the hottest region of the transformer winding, is one of the important factors that affects its insulating capacity [1–4]. TWHT that exceeds the allowable value will lead to damage of the transformer insulation. Therefore, it is of great significance to detect hotspot temperature anomalies in advance [5].

A lot of research has been done on the prediction of TWHT. A method of predicting the transformer top temperature based on the T-S fuzzy model was proposed [6]. However, the model had some limitations since top temperature is not actually the hotspot temperature. A TWHT prediction model was established based on the generalized regression neural network (GRNN) [7]. Additionally, the real-time estimation model for TWHT was established based on the Kalman filter algorithm, taking into consideration the load conversion and top oil temperature [8]. The least squares support vector machine (LS-SVM) and the genetic optimization support vector machine (GO-SVM) were also used to establish the transformer top temperature prediction model [9,10]. The above research achieved data tracking, but it is insufficient at predicting the TWHT anomaly in advance.

In recent years, the time series prediction model has been widely used in stock forecasting, traffic flow forecasting, and wind power forecasting [11]. However, it is rarely used in TWHT prediction. Scholars of the University of Queensland, Australia put forward the TWHT prediction model based on fuzzy information granulation (FIG) and the support vector machine (SVM) [12]. However, their study followed the stock forecasting model without redesigning the actual conditions of the TWHT prediction.

Since the study on the time series prediction and fluctuation range prediction of TWHT is insufficient, we propose a new prediction model for the fluctuation range of TWHT based on fuzzy information granulation and the chaotic particle swarm optimized wavelet neural network (FIG-CPSO-WNN). In this model, effective information is extracted from the original data using the FIG technique and is used to construct the wavelet neural network (WNN) prediction model. The structural parameters of the WNN prediction model are optimized using chaotic particle swarm optimization (CPSO). Additionally, the optimized model is eventually used to predict the fluctuation range of TWHT. The experiments and data analysis demonstrate that the model is more effective than other prediction models.

2. Fuzzy Information Granulation

Information granulation, which was first proposed by Professor Zadeh, is the process of dividing large amounts of complex information into multiple parts by certain rules, making each part have a relatively low computation complexity [13]. In this process, the original information is compressed with the extraction of effective information. Information granulation can be based on fuzzy set theory, rough set theory, and entropy space theory. In this paper, fuzzy set theory is used to study the fuzzy granulation process of TWHT time series. FIG can be divided to two procedures: (1) the division of granulation time window; and (2) the data fuzzification of each time window by certain rules [14].

The purpose of granulation time window is to decompose the TWHT time series into several subsequences as needed, in which each subsequence is a time window, while information fuzzification is the process of extracting effective information by establishing a reasonable fuzzy set for the data of each time window by certain rules [15].

A TWHT series $T = [t_1, t_2, \dots, t_n]$ is considered to be a time window in data fuzzification. Additionally, the objective is to establish a fuzzy particle P on the basis of T , that is, to determine the function:

$$P = A(t), t \in T. \quad (1)$$

The fuzzy particle P is mainly in the form of a triangle, a trapezoid, and an asymmetric Gaussian type. For the prediction of TWHT fluctuation range, we need to know the maximum and the minimum of the time window. Thus, the triangle type fuzzy particle is adopted, the subordinate function of which is:

$$A(t, LOW, R, UP) = \begin{cases} 0, & t < LOW; \\ \frac{t-LOW}{R-LOW}, & LOW \leq t \leq R; \\ \frac{UP-t}{UP-R}, & R < t \leq UP; \\ 0, & t > UP. \end{cases}, \quad (2)$$

where LOW , R , and UP are the variables obtained by fuzzy granulation of each time window. LOW is the minimum value of the data change in each time window. R and UP are respectively the average level and the maximum value of data changes in each time window.

The value of LOW , R , and UP are determined by the W.Pedrycz algorithm. The steps are outlined below.

- (1) Determine the mean value R . $R = \text{median}(T)$ where T is the time series put in ascending order.

- (2) Determine the lower bound LOW .

$$\text{Maximize } Q(LOW) = \frac{\sum_{t_k \leq R} A(t_k)}{R - LOW}. \quad (3)$$

- (3) Determine the upper bound UP .

$$\text{Maximize } Q(UP) = \frac{\sum_{t_k > R} A(t_k)}{UP - R}. \quad (4)$$

- (4) Determine the fuzzy particle P .

$$P = (LOW, R, UP). \quad (5)$$

LOW , R , and UP are the variables obtained by fuzzy granulation of each time window. LOW is the minimum value of the data change in each time window. R and UP are, respectively, the average level and the maximum value of data changes in each time window.

3. Chaotic Particle Swarm Optimized Wavelet Neural Network

Wavelet neural network (WNN) is a combination of wavelet analysis and neural network, which possesses both the time-frequency local property of wavelet transform and the non-linear mapping capability of neural network.

The topology of WNN includes the input layer, the hidden layer, and the output layer. In this paper, the WNN with single output is adopted. $\{x_1, \dots, x_i, \dots, x_n\}$ is the input of WNN as well as the TWHHT historical data after fuzzy granulation; y is the prediction output data of WNN; w_{ij} is the weight coefficient from the input layer to the hidden layer; w_j is the weight coefficient from the hidden layer to the output layer; ψ is the excitation function of the hidden layer nodes, and Morlet-base wavelet function is used here.

When the input of the i th node is x_i ($i = 1, 2, \dots, n$), the input to the j th hidden layer node is:

$$H_in(j) = \sum_{i=1}^n w_{ij}x_i. \quad (6)$$

The data of the hidden layer shall go through stretching transformation and translation transformation. Hence, the input value of the base wavelet function is:

$$W_in(j) = \frac{H_in(j) - b_j}{a_j}, \quad (7)$$

where a_j is the stretching factor and b_j is the translation factor. Considering the Morlet-base wavelet function.

$$\psi(x) = \cos(1.75x)e^{-\frac{x^2}{2}}, \quad (8)$$

the output of the hidden layer node can be calculated as:

$$H_out(j) = \cos(1.75 \times W_in(j))e^{-\frac{1}{2} \times (W_in(j))^2}. \quad (9)$$

The output node of WNN uses the Sigmoid function $f(x) = 1/(1 + e^{-x})$. So the final prediction result of the model can be expressed as:

$$y = f\left(\sum_{j=1}^m w_j H_out(j)\right). \quad (10)$$

And the prediction error of the model is defined as:

$$J = \sum_{p=1}^T |Y_p - y_p|, \quad (11)$$

where T is the number of training samples; Y_p and y_p are, respectively, the measured data and the predicted data for the P th sample.

The prediction error J depends on the weight coefficient w_{ij} and w_j , the stretching factor a_j , and the translation factor b_j . Therefore, it is necessary to select reasonable model parameters. The traditional WNN uses the gradient descent method to select parameters. However, the convergence is slow and it cannot effectively search for the global optimal result [16–18]. To solve this problem, this paper introduces the chaotic particle swarm optimization (CPSO).

Particle swarm optimization (PSO) has been widely used in many fields because of its easy implementation, fast convergence, and good robustness [19–21]. However, there are two problems in searching for the optimal solution with the standard PSO [22]:

- (1) The randomness of the initialization process may result in inferior solutions, affecting the convergence of the evolutionary process.
- (2) The solution obtained by this algorithm may be a local optimal solution rather than a global optimal solution.

In order to avoid the above problems, CPSO is introduced. It uses the chaotic series to initialize the position and velocity of particles, which maintains the population diversity and improves the searching ergodicity. Additionally, the current optimal position is taken to be the initial position to generate a new chaotic series. The solution set is renewed by replacing a particle of the original population with the optimal particle. It uses the randomness and ergodicity of chaotic motion to ensure the diversity of the population solutions and avoid a local optimal position.

In this paper, the chaotic series is generated by *Logistic* equation:

$$z_{n+1} = \mu z_n (1 - z_n), n = 0, 1, 2, \dots \quad (12)$$

When $\mu = 4$ and $0 < z_0 < 1$, it will generate a complete chaotic series. Therefore, any arbitrary initial value z_0 ($0 < z_0 < 1$) can generate a chaotic series (z_1, z_2, z_3, \dots) by iteration.

Based on the above analysis, to search for the optimal WNN structure parameters with CPSO, we should focus on the establishment of the fitness function and particle coding. In this paper, the prediction error of the WNN model is taken as the objective function, that is, the result of Equation (11) is taken as the fitness value. Real number encoding is used to encode the particles. The input weight w_{ij} , the output weight w_j , the stretching factor a_j , and the translation factor b_j , which are the parameters that need to be optimized, are taken as the position vectors of each particle, as shown in

$$X(M) = [w_{ij}, a_j, b_j, w_j], \quad (13)$$

where M is the number of particle swarms, i is the number of input layer nodes, and j is the number of hidden layer nodes. Therefore, the particle dimensionality $D = (i + 3) \times j$.

The algorithm flow using CPSO to train WNN is as follows:

- (1) Initialize the parameters to generate a random D -dimensional vector ($z_1 = (z_{11}, z_{12}, \dots, z_{1D})$) with the value of each component being between 0 and 1. Get N vectors z_1, z_2, \dots, z_n by iteration according to Equation (9). Additionally, transform the components of each vector into the corresponding value range.
- (2) Call the WNN to obtain the fitness value of the particles. Select M particles from the N populations to form the initial population. Additionally, generate the initial velocity of the M particles by chaotic series.

- (3) Update the individual extremum $pBest$ if the particle fitness value is superior to $pBest$. Additionally, update the global extremum $gBest$ if the particle fitness value is superior to $gBest$. Then update the particle's position and velocity.
- (4) Perform chaotic optimization to the current optimal position $P_{gBest} = (p_{g1}, p_{g2}, \dots, p_{gD})$. Additionally, map P_{gBest} to 0–1 according to:

$$y_i = \frac{p_{gi} - p_{\min}}{p_{\max} - p_{\min}}, \quad i = 1, 2, \dots, D. \quad (14)$$

Let $Y = [y_1, y_2, \dots, y_D]$. The chaotic series Y_1, Y_2, \dots, Y_m is generated after m times iterations by *Logistic* equation. The chaotic series is then mapped back to the original solution space according to:

$$P_{gj}^* = (p_{\max} - p_{\min}) \times Y_j + p_{\min}, \quad j = 1, 2, \dots, m. \quad (15)$$

So the new solution series is $P_{g1}^*, P_{g2}^*, \dots, P_{gm}^*$.

- (5) Calculate the fitness value of the new solution and replace the position of any particle in the current population with the best solution P^* .
- (6) Determine whether the termination condition is reached. If not, return to step (3) to continue iteration. Otherwise, the flow terminates and the selected parameters are obtained.

4. Prediction Model for Transformer Winding Hotspot Temperature Fluctuation Range

Based on the above studies, a new prediction model of TWHT fluctuation range based on fuzzy information granulation and chaotic particle swarm optimization wavelet neural network (FIG-CPSO-WNN) is proposed. The overall flow of the model is shown in Figure 1.

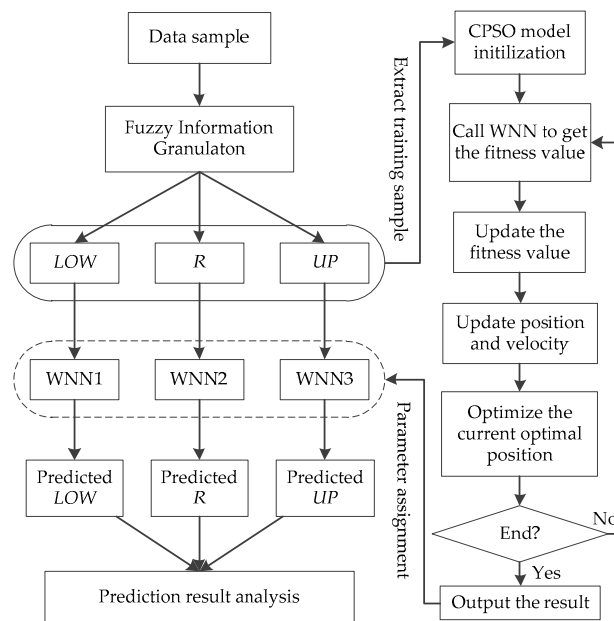


Figure 1. Flow chart of the new prediction model.

The modeling procedure is as follows:

- (1) Extract the sample data. Determine the size of granulation time window according to the sample data. The granulation data should be able to describe the trend of the original sample data. Perform FIG to sample data to obtain *LOW*, *R*, and *UP*.

- (2) If the amount of data is huge, the output layer node shall be determined according to forecast demand. The input layer node n shall be determined according to the intrinsic regularity of the historical data. Additionally, the number of hidden layer node shall be $m = 2n + 1$.
If the amount of data to be processed is limited, the output layer node shall be determined according to forecast demand, while the input layer node and the hidden layer node shall be determined by traversing method.
- (3) The structural parameters that need to be optimized are determined when the input layer, the hidden layer, and the output layer are determined. Encode the optimization object according to Equation (13), and take Equation (11) to be the fitness value function. Screen the structural parameters of the WNN by CPSO.
- (4) Use the designed WNN to predict LOW , R , and UP to obtain the fluctuation range of the TWHT.
- (5) Evaluate the performance of the model by the mean square error (MSE), the mean absolute error (MAE), and the correlation coefficient (r), calculated as:

$$MSE = \frac{1}{N} \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (16)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (17)$$

$$r = \frac{Cov(y, \hat{y})}{\sqrt{D(y)}\sqrt{D(\hat{y})}}. \quad (18)$$

5. Case Study

The analysis object of this paper is a three-winding, oil-immersed transformer equipped with optical fiber on-line temperature monitoring system. The parameters of the transformer are shown in Table 1. We recorded the data from 3 to 16 August 2016 with a sampling interval of half an hour.

Table 1. Transformer parameters.

Rated Parameters		Basic Structural Parameters	
rated capacity	24/24/16 MVA	oil weight	62,000 kg
rated voltage	220/115/37 kV	winding weight	39,482 kg
rated current	630/1205/2497 A	core weight	99,658 kg
rated frequency	50 Hz	tank length	10 m
connection type	YNyn0d11	tank width	2.5 m
cooling method	ONAN	tank height	4 m

5.1. Fuzzy Information Granulation

The transformer winding hotspot temperature recorded within the two weeks is shown in Figure 2. With continuous recording of 336 h and a sampling interval of half an hour, there are a total of 672 sample points. Four sample points, i.e., two hours, is taken as a time window to perform FIG, the result of which is shown in Figure 3.

In Figure 3, LOW , R , and UP are, respectively, the minimum, the approximate average, and the maximum temperature of each time window. The recorded temperature data is decomposed into the information particles LOW , R , and UP to describe the temperature range. This achieves the extraction of effective information and reduces the difficulty of data processing.

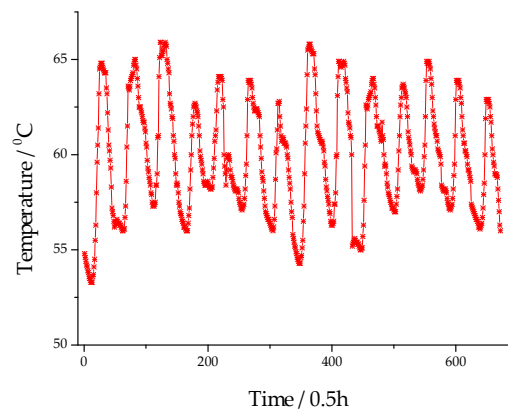


Figure 2. Recorded transformer winding hotspot temperature.

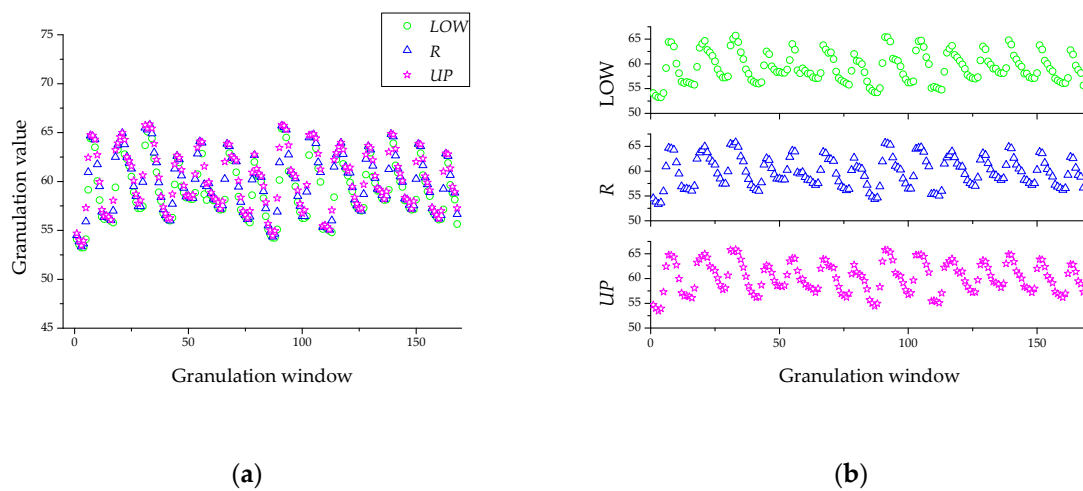


Figure 3. (a) General view of fuzzy information granulation (FIG); (b) separate views of FIG.

5.2. Establishment of Optimized WNN Prediction Model

In this paper, the structure and parameters of the WNN are selected based on R data, the approximate average temperature of each time window obtained by FIG. Additionally, the prediction model of TWHT fluctuation range based on CPSO-WNN is thereby established.

There are a total of 168 sample points of R data, which are divided into 14 groups. Group 1–7 is the training set and group 8–14 is the test set. For both the training set and the test set, the first six groups of data are taken as the input information, and the last group is taken as the output information.

It is necessary to consider the following three problems using WNN to establish the time series prediction model.

- (1) The selection of wavelet function.
- (2) The number of nodes for the input layer and the hidden layer.
- (3) The selection of the input weight w_{ij} , the output weight w_j , the translation factor b_j , and the stretching factor a_j .

Since the prediction result is influenced by the above three aspects, appropriate structure and parameters must be chosen for the model.

The prediction model is expected to forecast the fluctuation range of the TWHT one day in advance according to historical information. So, the number of output layer node is set to be 1. Only the number of nodes for the input layer and the hidden layer is to be determined. In this paper, a number of different structures of WNN are established with input layers of 1–6 and hidden layers of

3–15. To select a proper model, we analyze R data and compare the prediction accuracy. The number of iteration is set to be 100. The prediction accuracy is defined as the mean square error between the predicted value and the measured value. The result is shown in Figure 4.

As indicated by Figure 4, the prediction accuracy is highest when the numbers of input layer and hidden layer are, respectively, 1 and 6. Therefore, the “1-6-1” WNN structure is adopted. That means the data of group 6 and 7 are, respectively, the input and output for training. Additionally, the data of group 13 and 14 are, respectively, the input and output for testing. So, the TWHT fluctuation range of the next day can be predicted by entering the historical data of the previous day.

During the optimization of the input weight, output weight, translation factor, and stretching factor of the WNN by CPSO, the fitness value curve is obtained as shown in Figure 5. Additionally, the optimal parameters are shown in Table 2.

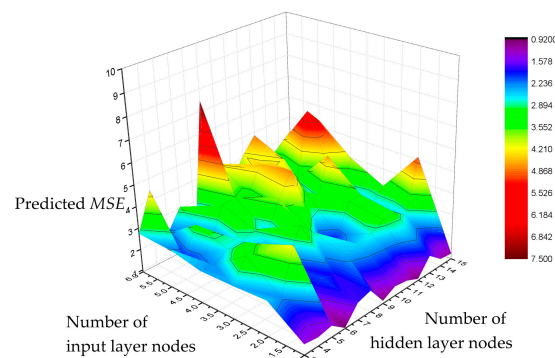


Figure 4. Comparison of prediction indicators.

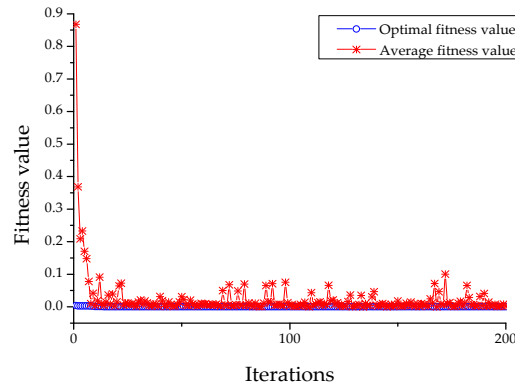


Figure 5. Fitness value curve.

Table 2. Parameter selection results.

Serial Number	w_{ij}	a_j	b_j	w_j
1	0.8199	0.3208	0.1225	0.3837
2	0.2874	0.4802	0.6329	0.6052
3	0.8830	0.2729	0.9809	0.5076
4	0.5415	0.5215	0.0292	0.3290
5	0.5623	0.2805	0.4745	0.6701
6	0.9254	0.2475	0.7413	0.4753

The parameters are assigned according to the selection result. Compared with the gradient descent method, the parameter selection method in this paper is more effective at improving the prediction accuracy of WNN.

5.3. Comparison of Four Prediction Models

To show the superiority of the FIG-CPSO-WNN time series prediction model, support vector regression (SVR) and Elman neural network model [23,24] are also studied in this paper. Four types of prediction models are established using the same sample data, which are, respectively, the SVR model, the Elman neural network model, the WNN model, and the FIG-CPSO-WNN model.

The error comparison of the prediction results of the four models is shown in Figure 6. The indicator takes the absolute value of the relative error.

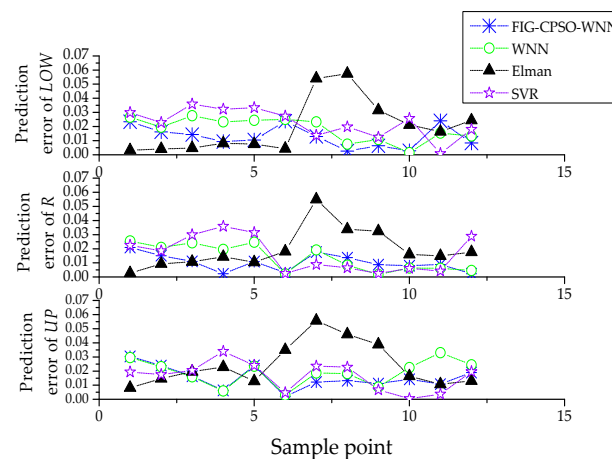


Figure 6. Comparison of prediction error.

It can be seen from Figure 6 that the FIG-CPSO-WNN model shows a more stable prediction performance. There is no prediction point with large error. However, the prediction for the training set can only demonstrate the fitting capacity of the model. It is essential to analyze the test set data in order to verify the model.

The predictions of the four models to the test set sample are shown in Figure 7. We can see that the prediction result of FIG-CPSO-WNN model is closer to the measured data than the other models.

Table 3 shows the quantitative performance evaluation of the four prediction models. A model is considered to be superior in performance if MSE and MAE are small and r is close to 1.

As can be seen from Table 3, the FIG-CPSO-WNN model is, overall, superior to the SVR model. It is superior to the Elman model in MSE and MAE . Additionally, it is overall superior to the WNN model, since the parameters of WNN are unreasonable. It is proved that it is effective to use the CPSO to select the parameters of WNN. The generalization ability and robustness of the model is improved.

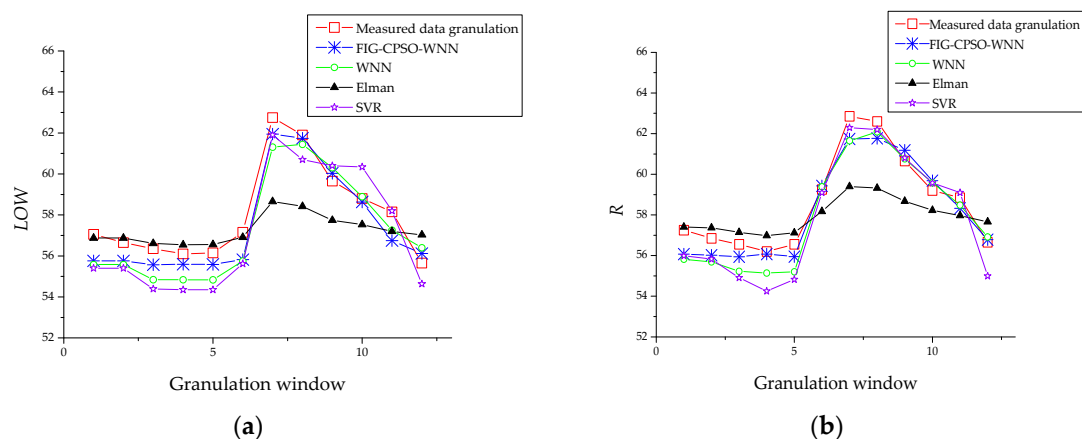
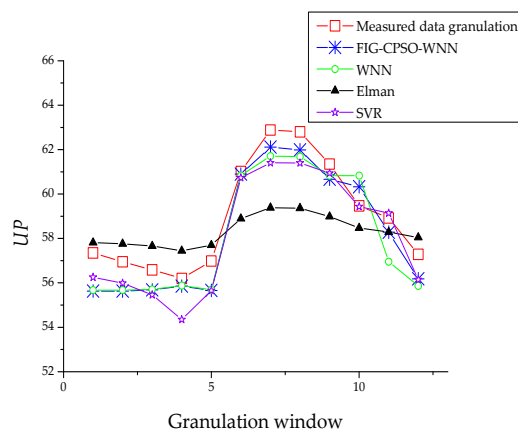


Figure 7. Cont.



(c)

Figure 7. (a) Comparison of *LOW* predictions; (b) comparison of *R* predictions; (c) comparison of *UP* predictions.

Table 3. Performance evaluation of three models.

Prediction Model		SVR	Elman	WNN	FIG-CPSO-WNN
<i>LOW</i>	<i>MSE</i>	1.3812	1.7679	1.1170	0.8385
	<i>MAE</i>	1.2786	1.2362	1.0265	0.7290
	<i>r</i>	0.9379	0.9708	0.9457	0.9681
<i>R</i>	<i>MSE</i>	1.1378	1.6338	0.9209	0.6830
	<i>MAE</i>	0.9269	1.2716	0.7743	0.5954
	<i>r</i>	0.9785	0.9863	0.9665	0.9689
<i>UP</i>	<i>MSE</i>	1.092	1.8314	1.2137	0.9779
	<i>MAE</i>	0.9432	1.5108	1.1000	0.8828
	<i>r</i>	0.9725	0.9946	0.9483	0.9729

5.4. Comparison of the Predicted Fluctuation Range with the Measured Data

Figure 8 shows the measured TWHT data and the predicted fluctuation range.

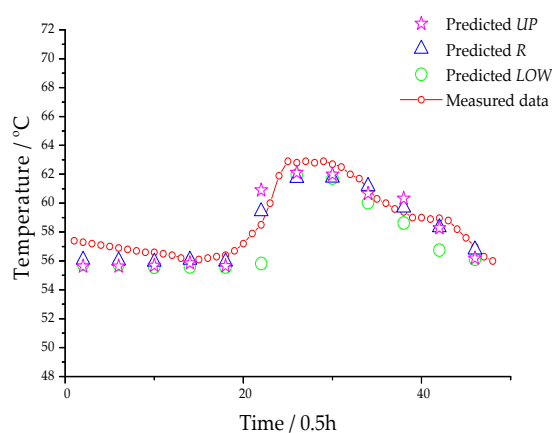


Figure 8. Comparison of predicted fluctuation range with the measured data.

It can be seen from Figure 8 that the results of this model are in good agreement with the actual data. Since the prediction error cannot be directly calculated, we choose the maximum, the mean, and the minimum of the actual data in each window to make comparison with the predicted range.

Using the data in Table 4, we obtain the prediction error shown in Table 5.

Table 4. Comparison of the measured data and the predicted data.

Time Window	Measured Data			Predicted Data		
	MIN	MEAN	MAX	LOW	R	UP
1	57.1	57.2	57.4	55.8	56.1	55.6
2	56.7	56.8	57	55.8	56	55.6
3	56.4	56.5	56.6	55.6	55.9	55.7
4	56.1	56.2	56.2	55.6	56.1	55.9
5	56.3	56.3	57.2	55.6	56	55.7
6	57.9	60	61.9	55.8	59.4	60.9
7	62.8	62.8	62.9	61.9	61.7	62.1
8	62	62.5	62.9	61.7	61.8	62
9	60	61	61.7	60	61.2	60.7
10	59	59.4	59.6	58.6	59.7	60.3
11	58.2	58.8	60	56.7	58.3	58.3
12	56	57	57.6	56.1	56.8	56.2

Table 5. Prediction error.

Evaluation Index	LOW	R	UP
MSE	0.9836	0.6318	1.1951
MAE	0.7917	0.5417	1.1167
r	0.9666	0.9823	0.9692

As can be seen from Table 5, the prediction error of this model is reasonable. So, the prediction results can provide reference for the operation and maintenance of transformers.

6. Conclusions

The transformer winding hotspot temperature has been rarely studied for time series prediction and fluctuation range prediction. Therefore, a new prediction model of the transformer winding hotspot temperature fluctuation range based on fuzzy information granulation and the chaotic particle swarm optimized wavelet neural network (FIG-CPSO-WNN) is proposed in this paper. The following conclusions are drawn from the experimental and analytical results.

- (1) Information granulation can extract useful information from the raw data, which reduces the complexity of target data. Additionally, the wavelet neural network (WNN) has a strong nonlinear mapping capability. In this paper, the two methods are combined to make effective predictions of the transformer winding hotspot temperature fluctuation range.
- (2) By designing the WNN according to the field data, we obtain a superior prediction performance compared with various prediction models. The feasibility of the FIG-CPSO-WNN model for predicting the transformer winding hotspot temperature fluctuation range is thereby demonstrated.
- (3) The proposed model has a high prediction accuracy and guiding significance to the operation and maintenance of transformers. The new model can not only be used in the prediction of transformer winding hotspot temperature fluctuation range, but also provides ideas for prediction modeling in other areas.

However, more in-depth research is needed to make the model more applicable to engineering practices. Issues to be studied include the selection of FIG window and the establishment of the model in other prediction areas.

Acknowledgments: This research was supported by State Grid Corporation Headquarters Science and Technology Project Funding (SGSG0000JSJS1700014) and the Fundamental Research Funds of Shandong University (Grant 2017JC012).

Author Contributions: Li Zhang, Wenfang Zhang, and Jinxin Liu designed the new prediction model for transformer winding hotspot temperature range, performed the data analysis, and prepared the manuscript. Tong Zhao, Liang Zou, and Xinghua Wang assisted the data acquisition, modeling, and performance analysis of the model. All authors discussed the results and approved the publication.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pradhan, M.K.; Ramu, T.S. Prediction of hottest spot temperature (HST) in power and station transformers. *IEEE Trans. Power Deliv.* **2004**, *18*, 1275–1283. [[CrossRef](#)]
2. McNutt, W.J.; Mciver, J.C.; Leibinger, G.E. Direct measurement of transformer winding hotspot temperature. *IEEE Trans. Power Appar. Syst.* **1984**, *103*, 1155–1162. [[CrossRef](#)]
3. Assuncao, T.; Silvino, J.L.; Resende, P. Transformer top-oil temperature modeling and simulation. *Trans. Eng. Comput. Technol.* **2006**, *15*, 240–245.
4. Taghikhani, M.A.; Gholami, A. Prediction of hottest spot temperature in power transformer windings with non-directed and directed oil-forced cooling. *Electr. Power Energy Syst.* **2009**, *31*, 356–364. [[CrossRef](#)]
5. International Electrotechnical Commission. *Power Transformers—Part 7: Loading Guide for Oil-Immersed Power Transformers*; IEC 60076-7; IEC: Geneva, Switzerland, 2005.
6. Chen, W.G.; Pan, C.; Yun, Y.X. Power transformer top-oil temperature model based on thermal-electric analogy theory. *Eur. Trans. Electr. Power* **2009**, *19*, 341–354.
7. Chen, W.G.; Xi, H.J.; Su, X.P.; Liu, W. Application of generalized regression neural network to transformer winding hot spot temperature forecasting. *High Volt. Eng.* **2012**, *38*, 16–21.
8. Su, X.P.; Chen, W.G.; Xi, H.J. Winding hotspot temperature prediction for oil-immersed power transformer based on Kalman filter algorithm. *High Volt. Eng.* **2012**, *38*, 1909–1916.
9. Keerthi, S.S.; Lin, C.J. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Comput.* **2003**, *15*, 1667–1689. [[CrossRef](#)] [[PubMed](#)]
10. Van Gestel, T.; Suykens, J.A.K.; Baestaens, D.E.; Lambrechts, A.; Lanckriet, G.; Vandaele, B.; De Moor, B.; Vandewalle, J. Financial time series prediction using least squares support vector machines within the evidence framework. *IEEE Trans. Neural Netw.* **2001**, *12*, 809–821. [[CrossRef](#)] [[PubMed](#)]
11. Chang, H.; Lee, Y.; Yoon, B. Dynamic near-term traffic flow prediction: System-oriented approach based on past experiences. *IET Intell. Transp. Syst.* **2012**, *6*, 292–305. [[CrossRef](#)]
12. Wang, H.; Hu, Z.J.; Zhang, M.L. A combined forecasting model for range of wind power fluctuation based on fuzzy information granulation and least squares support. *Trans. China Electrotechnol. Soc.* **2014**, *29*, 218–224.
13. Cui, Y.; Ma, H.; Tapan, S. Transformer hot spot temperature prediction using a hybrid algorithm of support vector regression and information granulation. In Proceedings of the Power and Energy Engineering Conference (APPEEC), Beijing, China, 12–14 April 2015.
14. Zadeh, L.A. Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets Syst.* **1997**, *90*, 111–127. [[CrossRef](#)]
15. Pawlak, Z. Granularity of knowledge, indiscernibility and rough sets. In Proceedings of the IEEE World Congress on Computational Intelligence, Piscataway, NJ, USA, 4–9 May 1998.
16. Zhang, Q.; Benveniste, A. Wavelet networks. *IEEE Trans. Neural Netw.* **1992**, *3*, 889–898. [[CrossRef](#)] [[PubMed](#)]
17. Amjady, N.; Keynia, F. A new neural network approach to short term load forecasting of electrical power systems. *Energies* **2011**, *4*, 488–503. [[CrossRef](#)]
18. Maniezzo, V. Genetic evolution of the topology and weight distribution of neural networks. *IEEE Trans. Neural Netw.* **1994**, *5*, 39–53. [[CrossRef](#)] [[PubMed](#)]
19. Bargiela, A.; Pedrycz, W. *Granular Computing: An Introduction*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003.
20. Kennedy, J.; Eberhart, R.C. Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; pp. 1942–1948.
21. Chen, Z.; Xiong, R.; Wang, K.; Jiao, B. Optimal energy management strategy of a plug-in hybrid electric vehicle based on a particle swarm optimization algorithm. *Energies* **2015**, *8*, 3661–3678. [[CrossRef](#)]

22. Khanesar, M.A.; Teshnehlab, M.; Shoorehdeli, M.A. A novel binary particle swarm optimization. In Proceedings of the IEEE Conference on Control & Automation, Mediterranean, Athens, Greece, 27–29 June 2007.
23. Swift, G.; Molinski, T.S.; Lehn, W. A fundamental approach to transformer thermal modeling—Part I: Theory and equivalent circuit. *IEEE Trans. Power Deliv.* **2001**, *16*, 171–175. [[CrossRef](#)]
24. Lesieutre, B.C.; Hagman, W.H.; Kirtley, J.L. An improved transformer top oil temperature model for use in an on-line monitoring and diagnostic system. *IEEE Trans. Power Deliv.* **1997**, *12*, 249–256. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).