Article

Research on Transformer Condition Prediction Based on Gas Prediction and Fault Diagnosis

Can Ding [ID], Wenhui Chen, Donghai Yu * and Yongcan Yan

College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443000, China; dingcan@ctgu.edu.cn (C.D.); chenwenhui@ctgu.edu.cn (W.C.); yanyongcan@ctgu.edu.cn (Y.Y.)
* Correspondence: yudonghai@ctgu.edu.cn

Abstract: As an indispensable part of the power system, transformers need to be continuously monitored to detect anomalies or faults in a timely manner to avoid serious damage to the power grid and society. This article proposes a combined model for transformer state prediction, which integrates gas concentration prediction and fault diagnosis models. First, based on the historical monitoring data, each characteristic gas sequence is subjected to one optimal variational mode decomposition (OVMD) and one complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN). The decomposed sub-sequences are input into a bi-directional long short-term memory network (Bi-LSTM) optimized by the sparrow search algorithm (SSA) for prediction, and the predicted value of each sub-sequence was then superimposed to be the predicted value of the characteristic gas. We input the predicted values of each gas into the improved sparrow search algorithm-optimized support vector machine (ISSA-SVM) model, which can output the final fault type. After the construction of the combined model of state prediction is completed, this paper uses three actual cases to test the model, and at the same time, it uses the traditional fault diagnosis methods to judge the cases and compare these methods with the model in this paper. The results show that the combined model of transformer state prediction constructed in this paper is able to predict the type of transformer faults in the future effectively, and it is of great significance for the practical application of transformer fault type diagnosis.

Keywords: condition prediction; dissolved gas; fault diagnosis; gas prediction; oil-immersed transformer

1. Introduction

As a vital piece of equipment in the power system, the normal operation of transformers is crucial to the stability and reliability of the power supply. However, there is a potential risk of transformer faults due to power grid load fluctuations, changing environmental conditions, and aging equipment. Therefore, it is necessary to predict the transformer condition and take timely maintenance measures. Condition prediction of transformers is conducive to improving the reliability and stability of power supply [1], optimizing the maintenance strategy; realizing effective management of energy, energy saving, and emission reduction; and providing technical support and guarantees for the construction of the smart grid and sustainable energy development.

Condition prediction of oil-immersed transformers is divided into gas prediction and fault diagnosis based on gas prediction [2,3]. For the prediction of dissolved gases in oil, the early methods for predicting gas concentration include gray prediction [4] and the autoregressive integrated moving average (ARIMA) model [5], etc. These methods struggle to capture the characteristic regularity among the time series data in depth; moreover, they usually have a bad performance when facing a complex model, so there is a certain limitation imposed on the accuracy of gas concentration prediction. With the rapid development of intelligent technology, machine learning and deep learning models have been widely used in time series data prediction, along with random forests [6], support vector

machine regression [7,8], BP neural networks (BPNN) [9], and recurrent neural networks (RNN) [10], etc. After application, there is a substantial improvement in accuracy over traditional prediction methods. For the fault diagnosis part of the transformer, dissolved gas analysis (DGA) in oil [11,12] is still the main detection method at present, based on which two major categories of traditional detection methods and machine learning methods have been developed. Traditional detection methods currently contain the IEC three-ratio method [13], Duval's triangle law [14], and Rogers' ratio method [15]. These methods have a high fault diagnosis rate under certain conditions and have a certain reference, but due to the complexity of the actual environment in which the transformer is located, they usually lead to a decrease in the accuracy of fault prediction, so it is necessary to improve the fault diagnosis rate with other effective means. With the popularity of intelligent machine learning, relying on a large amount of gas concentration data, these intelligent algorithms can fully explore the hidden information between various types of gas concentration data and the operating state in which the transformer is located, which makes up for the shortcomings of the traditional diagnostic methods. Zhikai Xing [16] used the ARIMA model to predict dissolved gases and used the boost marine predators algorithm (BMPA) to optimize the parameters in the ARIMA algorithm, and their results showed that the prediction model has a high accuracy in predicting the operating status of the transformer. Wenqian Zhang [17] proposed to decompose the dissolved gas sequence into multiple sub-sequences using the CEEMDAN decomposition method. Each sub-sequence is individually predicted and superimposed using a temporal convolutional network (TCN). The final gas concentration prediction results exhibited higher prediction accuracy than other intelligent prediction algorithms. Xinbo Huang [18] and others constructed a hybrid kernel limit learning machine and optimized the parameters of the hybrid kernel function using the gray wolf optimization algorithm (GWO) for the fault diagnosis of transformers, and the results show that compared with BP neural network, the model proposed in their paper improves the accuracy of fault diagnosis and has a better generalization performance. To address the issue of imbalanced transformer fault type data, Suchandan Das [19] first utilized SMOTE (synthetic minority oversampling technique) to balance the fault type data. Next, they used recursive feature elimination (RFE) to optimize the gas features. Lastly, three classification algorithms were used to validate the proposed model. The test results showed that the diagnostic accuracy reached 97.43% in the IEC TC-10 database.
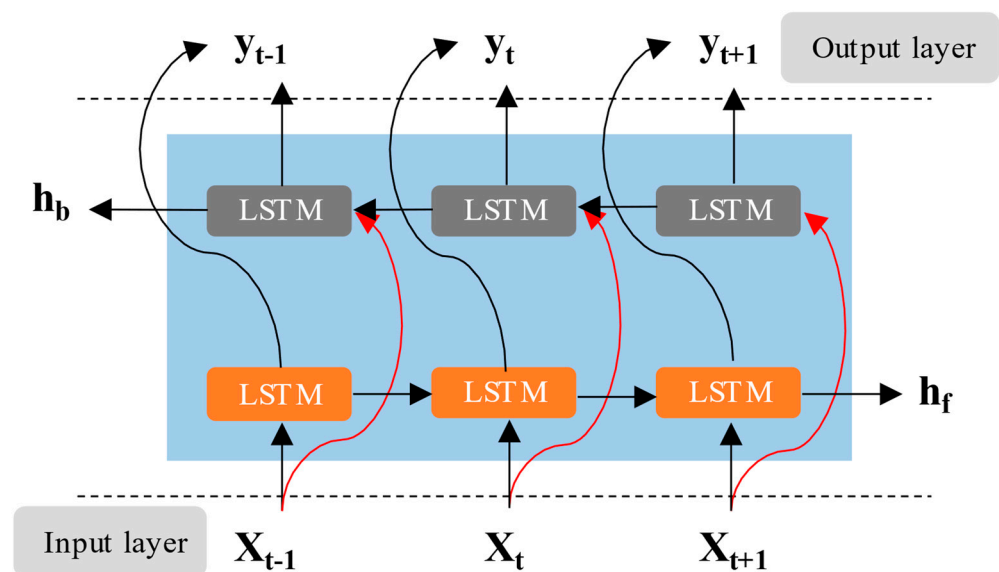
Currently, there have been numerous studies on predicting transformer gas concentrations and diagnosing faults. However, there are relatively few studies that combine both fields to predict the future state of transformers. This paper proposes a combined model for transformer state prediction, which is composed of a dissolved gas concentration prediction model and a fault diagnosis model based on dissolved gas prediction. In the gas concentration prediction model, this paper uses the optimal variational mode decomposition (OVMD) and the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) to decompose the gas concentration data; the subsequence after secondary decomposition excludes as much information as possible from the original time series data, which are useless and affect the prediction results. Then, a bi-directional long and short-term memory network (Bi-LSTM) [20–22], optimized by the sparrow search algorithm (SSA), was used to predict each subsequence, and the predicted value of each subsequence was superimposed as the predicted value of the characteristic gas concentration. Subsequently, the predicted values of each feature gas concentration are used as input. The transformer fault diagnosis model will determine whether a fault will occur in the transformer at a future time based on the relationship between the changes in gas concentration data and the fault types. The model will also output the corresponding fault type. This fault diagnosis model utilizes the support vector machine (SVM) classification algorithm. Prior to this, the diagnostic model has been trained and tested on a large amount of actual fault data. An improved sparrow search algorithm (ISSA) is employed for hyperparameter optimization to obtain the optimal values of the crucial C and gamma parameters in SVM. In order to verify the excellent performance of the state prediction

model, three real cases are chosen for the testing of this model. The results indicate that the model accurately predicted the fault types of all three cases, demonstrating the excellent performance of the proposed model in predicting future faults in transformers.

2. Introduction to Combinatorial Models

2.1. Gas Concentration Prediction Model

For the gas concentration prediction model, Bi-LSTM is selected as the prediction algorithm for processing time series data in this paper. The bidirectional long short-term memory network makes up for the defect in LSTM that can only learn unidirectionally and reads from the front-to-back and back-to-front, respectively, when processing the data, and the final output are the data after the combination of the two. In this way, the Bi-LSTM model not only learns the state at the current moment but also obtains more comprehensive information from the context; therefore, the data can be better modeled and predicted. The Bi-LSTM network model is shown in Figure 1.
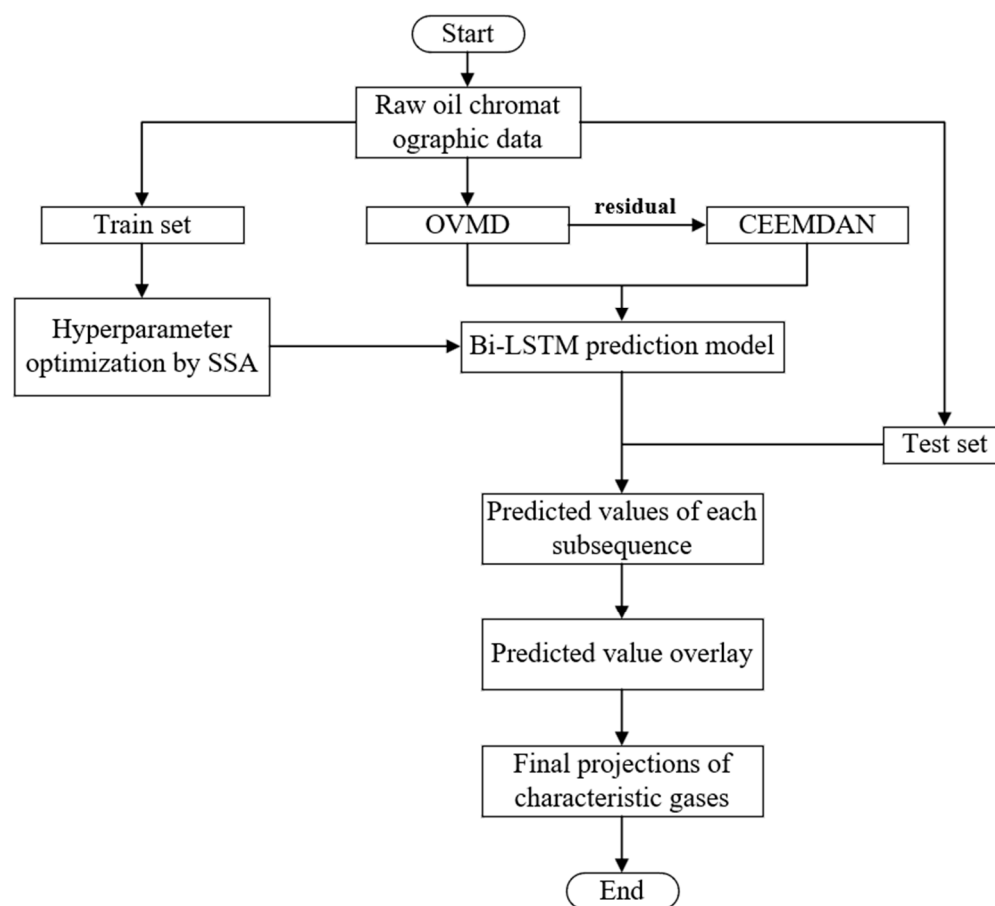

Figure 1. Principle of the Bi-LSTM model.

In Figure 1, h_f represents the output values of the hidden layer at each time step on the forward channel, while h_b represents the output values of the backward hidden layer at each time step.

Directly using Bi-LSTM to predict the original gas concentration data will be affected by the noise and harmonic components contained in the data; in view of this, it is proposed that the gas concentration data be decomposed via VMD and CEEMDAN quadratic decomposition so as to facilitate the full exploration of the potential links between the data. VMD is a non-recursive variational mode decomposition method for signal processing, commonly used to handle nonlinear time series. Its decomposition process primarily involves transforming the constrained variational problem into an unconstrained problem by introducing penalty factors and Lagrange operators, followed by iterative updates to obtain the modal components. VMD has good noise immunity and can overcome the issue of frequency aliasing. CEEMDAN is a type of complete ensemble empirical mode decomposition (EEMD), which adds white noise during the decomposition process to obtain the intrinsic mode functions (IMF) of each order after EMD decomposition, effectively reducing reconstruction errors during the decomposition phase.

First, the original gas data are subjected to a variational mode decomposition (VMD) to generate a series of sub-sequences with more pronounced trends, which are easier for prediction. The optimal number of decomposition levels is determined using the central frequency method. Next, for the residual sequences with large fluctuations and

irregularities resulting from the VMD decomposition of the original gas data, a secondary decomposition is conducted using the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) method to create further multiple regular components. Before using Bi-LSTM to predict the decomposed subsequence, it is necessary to optimize the parameters of the prediction algorithm. In this paper, the optimization of Bi-LSTM parameters utilizes the sparrow search algorithm. The parameters to optimize include the learning rate, number of training iterations, and the nodes in two hidden layers. By employing the optimization algorithm to individually optimize parameters for each sub-sequence's training set, the optimal parameters are applied to the Bi-LSTM to construct the SSA-BiLSTM model. Subsequently, this model is used to test the testing sets of the sub-sequences, yielding high-accuracy values. Finally, the superposition of the final sub-sequence values is the predicted concentration of the characteristic gas. It should be noted that the Bi-LSTM needs to be optimized with the sparrow search algorithm before predicting each subsequence of each characteristic gas. The flowchart of the whole gas concentration prediction model is shown in Figure 2.

Figure 2. Flowchart of the gas concentration prediction model.

2.2. Fault Diagnosis Model

The fault diagnosis part of the transformer outputs one of the multiple fault types, which is a classification model, and in this paper, the support vector machine [23] is used to classify the fault types for judgment. The support vector machine used kernel function widely for radial basis function (RBF). When the RBF kernel function is used, two important hyperparameters in the support vector machine need to be adjusted: the kernel function coefficient gamma and the penalty coefficient C. The classification effect of the final support vector machine model is closely related to the selection of these two hyperparameters. In this paper, the ISSA [24–26] is chosen to optimize the hyper-parameters of SVM, which has

the characteristics of fast optimization speed and avoiding falling into local optimum [27]. Since the input feature dimension is too low when the support vector machine is trained on the data, it will be difficult for SVM to explore the laws in it fully; according to the conclusion of the previous research, the fault type of the transformer is not very correlated with the content of CO and CO_2 gases, so it is not taken as the input feature gas for training. As for the original input features, there are only five kinds of gases, H_2, CH_4, C_2H_6, C_2H_4, and C_2H_2, which represent fewer training features for the support vector machine; in order to ensure the accuracy of fault classification, the training data need to be extended dimensionally during training [28]. According to ref. [28], in addition to the five original characteristic gases, new input features are required and are shown in Table 1.

Table 1. New input features.

Num	Feature	Num	Feature
1	CH_4/H_2	7	$CH_4/(C_1 + C_2)$
2	C_2H_2/C_2H_4	8	$C_2H_6/(C_1 + C_2)$
3	C_2H_4/C_2H_6	9	$(CH_4 + C_2H_4)/(C_1 + C_2)$
4	$C_2H_2/(C_1 + C_2)$	10	$C_1 + C_2$
5	$H_2/(H_2 + C_1 + C_2)$	11	$C_1 + C_2 + H_2$
6	$C_2H_4/(C_1 + C_2)$	12	$CH_4 + C_2H_4$

In Table 1, C_1 represents first-order hydrocarbons, which refers to CH_4 in the text; similarly, C_2 represents second-order hydrocarbons, which are the sum of the gas concentrations of C_2H_2, C_2H_4, and C_2H_6.

After the processing of the data, the ISSA is able to find the hidden connection between the input information and the fault type among the multidimensional data. The hyperparameters after optimization are applied to SVM to compose the ISSA-SVM fault diagnosis model. The specific process steps of the ISSA-SVM fault diagnosis model are as follows:

1. Due to the fact that the values of the input feature data are too different in size, in order to save computation time and improve the training efficiency of the model, it is necessary to normalize the size of the data so that all the values are within the range of [0,1]. The processing method is shown as in Equation (1):
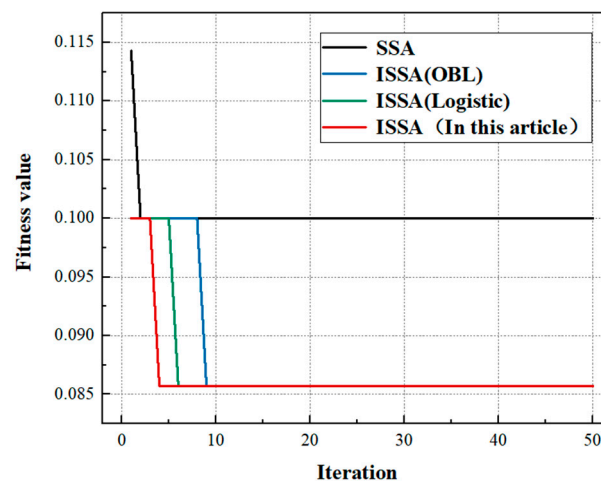
$$X_N = \frac{X - X_{min}}{X_{max} - X_{min}}. \tag{1}$$

In Equation (1), X_{min} and X_{max}, respectively, represent the minimum and maximum values of the data before normalization.

2. We randomly divide the dissolved gas data, corresponding to various fault types, into training sets and testing sets, use the training set and the extended data to find the optimal hyperparameters for the ISSA-SVM model, and apply the hyperparameters optimized by ISSA to the SVM classifier.

3. Using the optimized SVM fault diagnosis model to diagnose faults in the test set, after thorough learning and fitting on the training set, SVM can fully explore the potential correlations between different types of faults and data in numerous dimensions and ultimately output the fault diagnosis results.
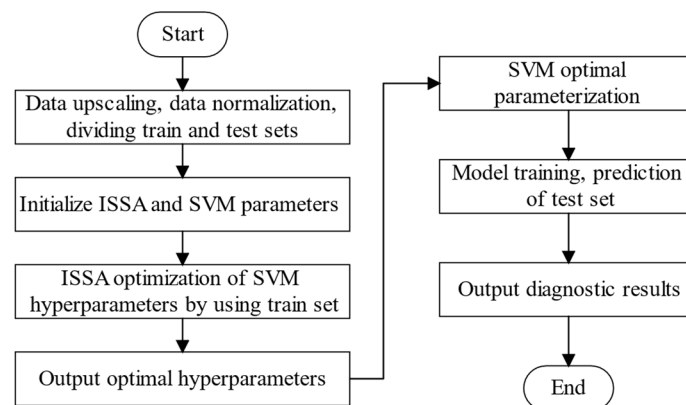
In ref. [28], the author of this paper made improvements to the original SSA algorithm, which mainly include initializing the population with Sin chaotic sequences, introducing a dynamic adaptive factor into the position update formula of the SSA algorithm, incorporating a reverse learning strategy, and applying Cauchy mutation. These methods help the SSA algorithm to escape local optima and improve convergence speed. Detailed explanations of these improvements can be found in ref. [28], and the present paper will not elaborate further. At the same time, this paper utilizes a set of actual transformer fault type data; after training and optimizing the training set obtained from this data, a set of optimal

SVM hyperparameter values was obtained. The SVM model with these hyperparameters achieved an accuracy of 91.43% on the test set, significantly higher than KNN (k-nearest neighbors), DELM (deep extreme learning machine), and other classification algorithms. Additionally, to verify the effectiveness of the improved method, the fitness curves obtained from two improved methods, one using the opposite learning strategy (OBL) and the other employing logistic regression, are compared. The results are shown in Figure 3. It can be observed that the improved method proposed in this paper exhibits a faster convergence speed compared to the other improved methods, demonstrating the excellent performance of the proposed approach. Therefore, the fault diagnosis models used in the three cases of this paper all adopt the ISSA-SVM model with the above-mentioned hyperparameters.

Figure 3. Fitness curves of the three ISSA algorithms.

The specific flowchart of the ISSA-SVM fault diagnosis model is shown in Figure 4.

Figure 4. Flowchart of the ISSA-SVM fault diagnosis model.

2.3. Combinatorial Model

Condition prediction technology refers to the prediction of the future operating state via the regular mining of known information on the basis of the currently observed and collected data. Most of the current research on the state prediction of oil-immersed transformers is also being carried out based on the DGA technique. The future concentration values predicted in the gas concentration prediction model are input to the fault diagnosis model for classification and diagnosis, which can realize the prediction of faults in advance. Compared with the simple fault diagnosis algorithm, fault prediction can detect and warn of the occurrence of faults in advance, thus helping operation and maintenance personnel to take relevant measures in time to reduce or avoid losses. The combined transformer

state prediction model constructed in this paper combines the gas concentration prediction model and the fault diagnosis model to predict the future operating state of oil-immersed transformers. The complete flow of the combined model for transformer state prediction is as follows:

1. Collect the dissolved gas concentration data to be predicted and also collect the data of the transformer under different faults and normal states; divide the gas concentration data into training set and test set data and carry out two different modal decompositions of the time series data to decompose them into multiple sub-sequences with a certain degree of regularity, which are easy to predict.

2. Use the sparrow search algorithm to optimize the parameters of the two-way long and short-term memory network, predict the multiple sub-sequences of the test set, and then superimpose the output values of the sub-sequences and reconstruct them into the predicted values of the final dissolved gas concentration. The same method is used to decompose, predict, and reconstruct different characteristic gases and output the final predicted value.

3. Normalize and preprocess the historical fault data of the transformer and input the data of the characteristic gases into the SVM optimized by the improved sparrow search algorithm after dimensional upgrading; the SVM will mine the correlation among them and train the diagnosis on the historical fault data and fault types.

4. The final predicted value data in step (2) are used as input to carry out the diagnosis process in step (3), and the trained ISSA-SVM model will evaluate the input data accordingly and output whether there is a fault in the transformer in the state of the predicted value, as well as the corresponding fault type.

The flowchart of state prediction based on gas concentration prediction and the transformer fault diagnosis model is shown in Figure 5. In Figure 5, the seven fault diagnosis results at the bottom are represented from left to right as low-temperature overheating, medium-temperature overheating, high-temperature overheating, normal state, partial discharge, low-energy discharge, and high-energy discharge.
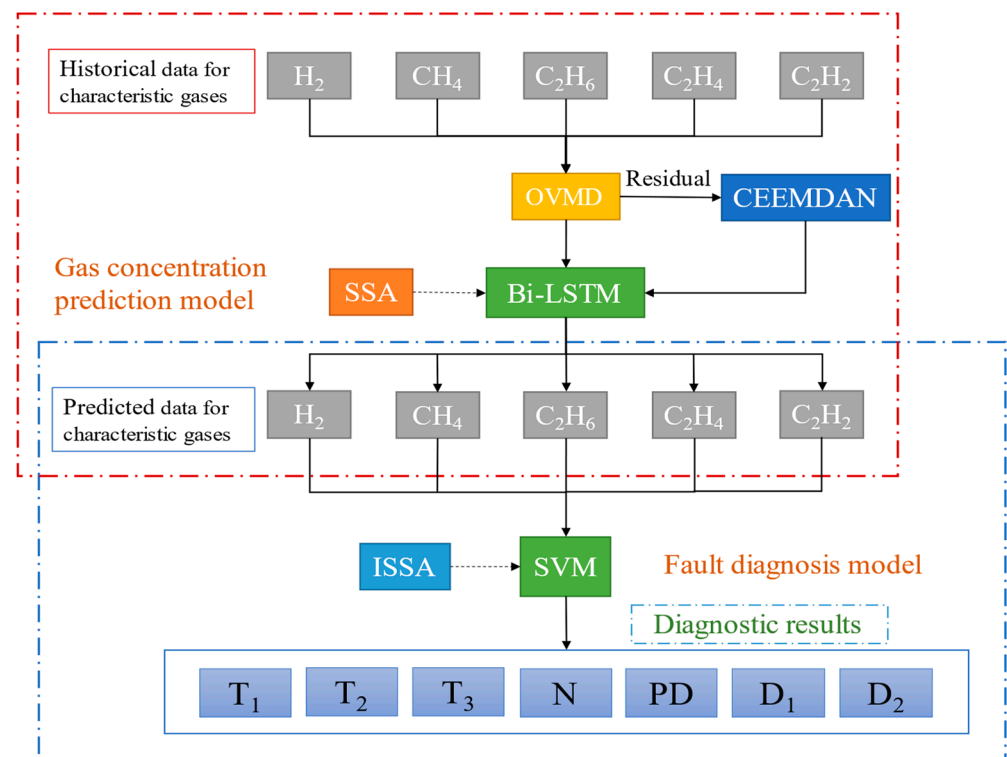

Figure 5. Oil-immersed transformer condition prediction flowchart.

3. Case Studies

In this chapter, three actual cases of transformer faults are selected to judge and analyze the performance of the state prediction combination model proposed in this paper, and the implementation process of the model is specifically introduced.

3.1. Case One

In this case, an oil-immersed transformer was selected that had previously faulted, and the corresponding historical oil chromatographic data are shown in Table 2. The transformer failed on 31 July, and the failure was diagnosed as a medium-temperature overheating fault on site. Prior to this, the sampling time for this transformer was once in half a month.

Table 2. Transformer oil chromatographic data (μL/L).

Date	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
1 March	6.3	4.9	5.8	3.4	0
16 March	11.5	24.5	6.3	16.2	0
31 March	26.8	40.2	35.7	56.1	0.4
15 April	32.7	61	41.3	134.9	0.4
1 May	51.8	44.6	56.9	143.5	0.6
16 May	46.6	135.2	48.2	172.4	0.5
31 May	48.6	120.1	98.8	229.6	0.8
15 June	52	176.3	103.1	202.5	1.2
1 July	49.6	157.3	92	293.1	1.1
16 July	67.2	199.1	123.9	335.8	1.2
31 July	224.9	152.6	371.4	71.2	1.4

For oil chromatography data prediction, three adjacent sampling time points are used to predict the value at the next time point; for example, the data on 1 March, 16 March, and 31 March are used as inputs to predict the gas concentration on 15 April, and so on. The data information before 31 July are used as the training set, the data on 31 July are used as the test set, the subsequence of the training set data after the quadratic modal decomposition is put into the SSA-BiLSTM model as the input for training, respectively, and parameter optimization is carried out with the SSA algorithm. After the training is completed, the Bi-LSTM model is used to predict the gas concentration values on 31 July based on the data from the first three instances of failure; then, the prediction performance of the model is judged based on the actual values.

The methane concentration prediction is now specifically analyzed as an example. The methane gas sequence is first subjected to the OVMD step, with the basic parameters for VMD set as follows: the constraint balancing parameter alpha is approximately 2500; the time step tau is 0.3; DC = 0 indicates no direct current component; the initial angular frequency change parameter is set to init = 1; and the convergence criterion tolerance is set to tol = 1×10^{-7}. In order to prevent the phenomenon of center frequency being too close, the optimal number of decomposition layers selected for the prediction of methane gas is six layers. At this point, the methane sequence is decomposed into six sub-sequences and one residual sequence. The residual sequence is then decomposed into three sub-sequences using CEEMDAN. Next, the nine sub-sequences are divided into training and testing sets and normalized. Due to the small amount of data in the case, data before July 31 are used as the training set, the data on July 31 are used as the testing set. The training set portions of the aforementioned nine sub-sequences are optimized for hyperparameters using the SSA-BiLSTM model. The objective hyperparameters for optimization include the learning rate, the number of training iterations, and the node count in two hidden layers, each with specific constraints on their optimization ranges. The population size of the sparrow search algorithm is 10, the number of iterations is 100, the ST of the forewarners is 0.8, and the PD ratio of the discoverers is 0.7. The optimal hyperparameters of the Bi-LSTM prediction model for each subsequence after optimization are shown in Table 3.

Table 3. The values of the hyperparameters for each subsequence take.

	Learning Rate	Training Times	Hidden Layer 1	Hidden Layer 2
VMF1	0.00911	89	95	93
VMF2	0.00547	96	64	71
VMF3	0.00866	84	88	74
VMF4	0.00828	80	97	47
VMF5	0.00536	98	14	96
VMF6	0.00703	15	52	88
IMF1	0.00723	51	7	58
IMF2	0.00415	99	85	32
IMF3	0.00363	90	93	51

In Table 3, VMF means the subsequences generated by the optimal variational mode decomposition of the original gas data, and IMF means the subsequences generated by the secondary decomposition of the residual sequence by the CEEMDAN algorithm. When predicting each subsequence, the hyperparameters corresponding to the nine subsequences were applied to the Bi-LSTM model, respectively, while the methane concentration values on the day of 31 July were predicted based on the sampling data of 15 June, 1 July, and 16 July. Figure 6 shows the prediction of the Bi-LSTM model for nine subsequences of methane data on 31 July. The predicted values of each subsequence are superimposed to obtain the predicted concentration value of methane on 31 July. Similarly, the same steps are taken for the other characteristic gases besides methane, and their predicted values on 31 July are shown in Figures 7–10. The concentration predicted values and actual data for the five characteristic gases on 31 July are shown in Table 4.

As can be seen from the comparison between the predicted value and the true value in Table 4, the results predicted by the SSA-BiLSTM model with quadratic modal decomposition are very close to the true value, which, to a certain extent, indicates that the prediction model has a good performance and provides a more accurate concentration value for the subsequent SVM fault diagnosis, which is conducive to the correct prediction for the state of the transformer.

The predicted values of each gas in Table 4 on 31 July are inputted into the ISSA-SVM fault diagnosis model, and the result predicted by the diagnosis model is the medium-temperature overheating fault, which is the same as the actual fault type of the transformer. Then, the traditional diagnosis method is used to judge the results; in this case, Duval's triangle law and the three-ratio method are chosen, the percentage contents of CH_4, C_2H_4, and C_2H_2 derived from Table 4 are inputted into Duval's triangle model, and the diagnostic result of the model is shown in Figure 11. The type of the fault falls in the red region, and the diagnostic fault is the high-temperature overheating fault, which is different from the actual situation. In the three-ratio method, the ratio of C_2H_2/C_2H_4, CH_4/H_2, and C_2H_4/C_2H_6 is calculated, respectively; then, the corresponding code is obtained by checking the table. Through the combination of the code, the type of fault can be judged. In this case, we know that the code is 021 through the existing table, and the diagnosis result is the medium-temperature overheating fault, which is the same as the real situation.
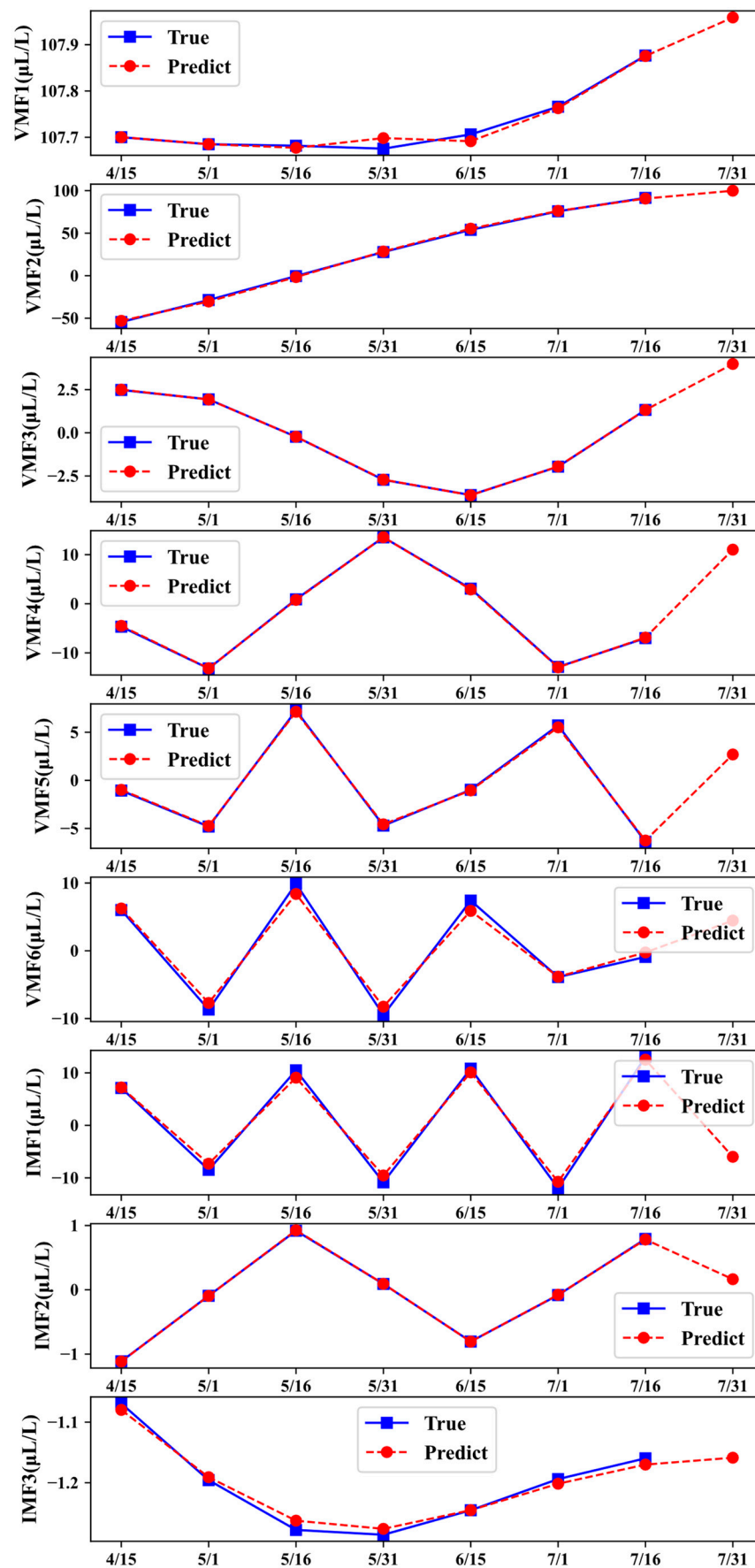
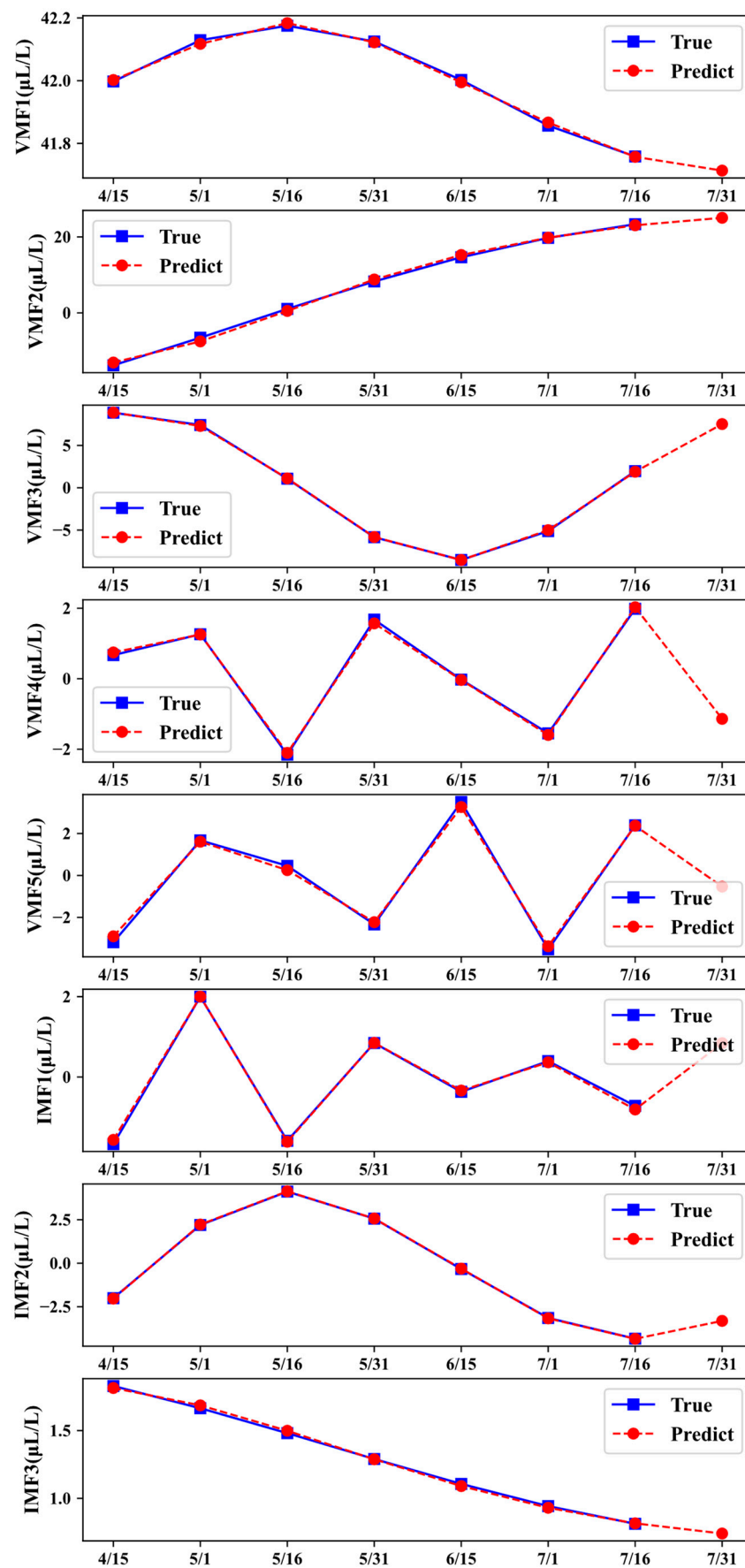**Figure 6.** Predicted results of methane gas for each subsequence on 31 July.

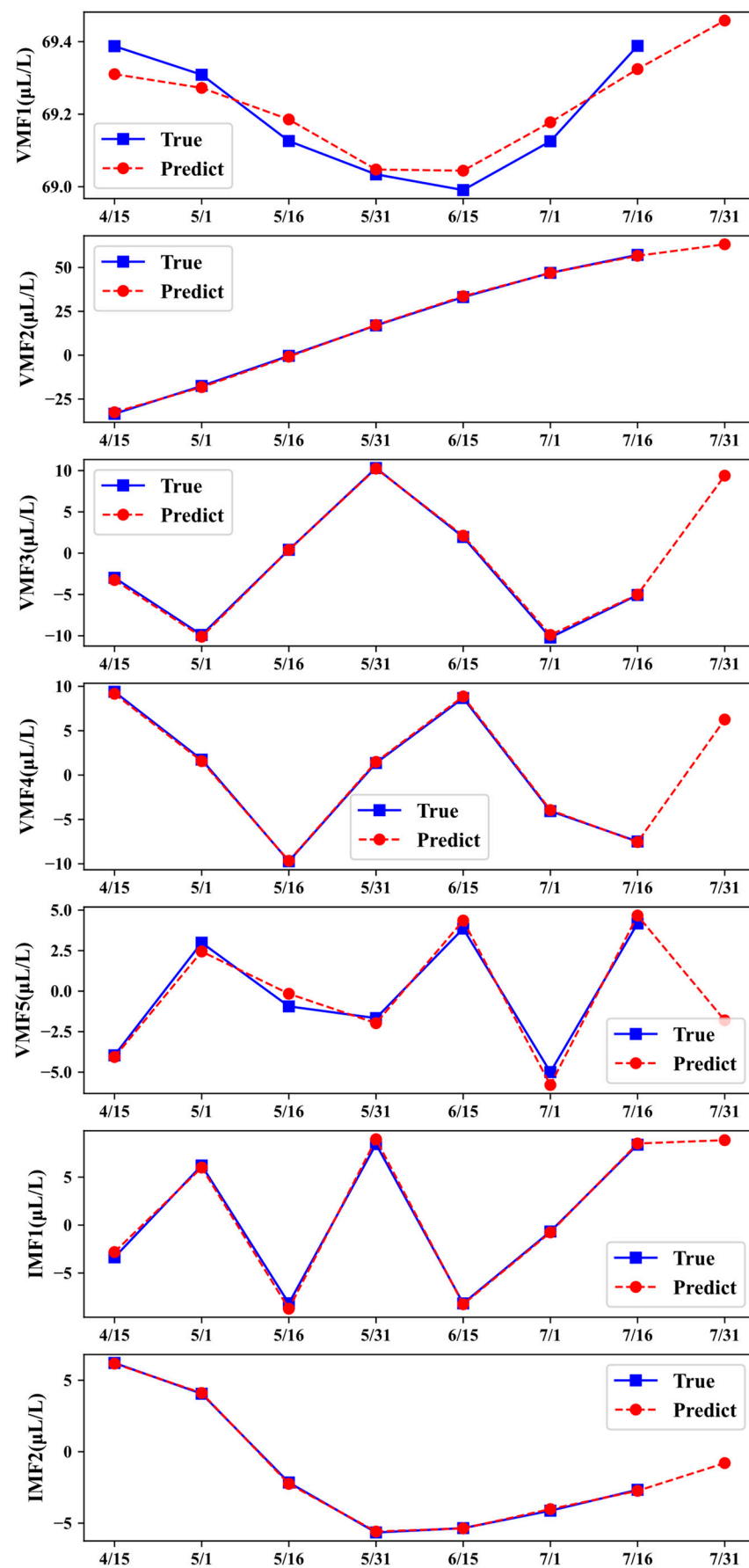**Figure 7.** The prediction results for the hydrogen subsequence on 31 July.

Figure 8. The prediction results for the ethane subsequence on 31 July.
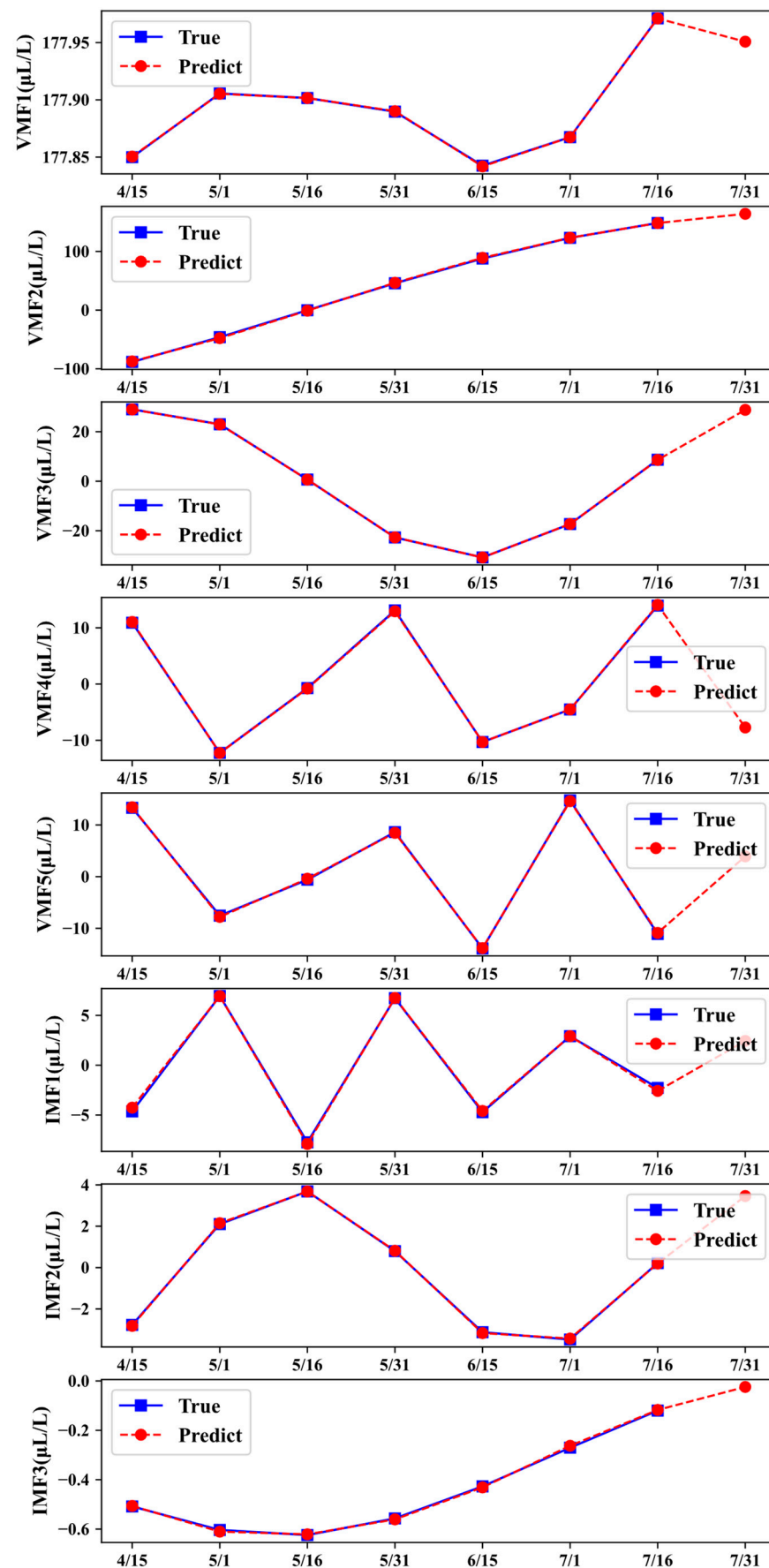
Figure 9. The prediction results for the ethylene subsequence on 31 July.
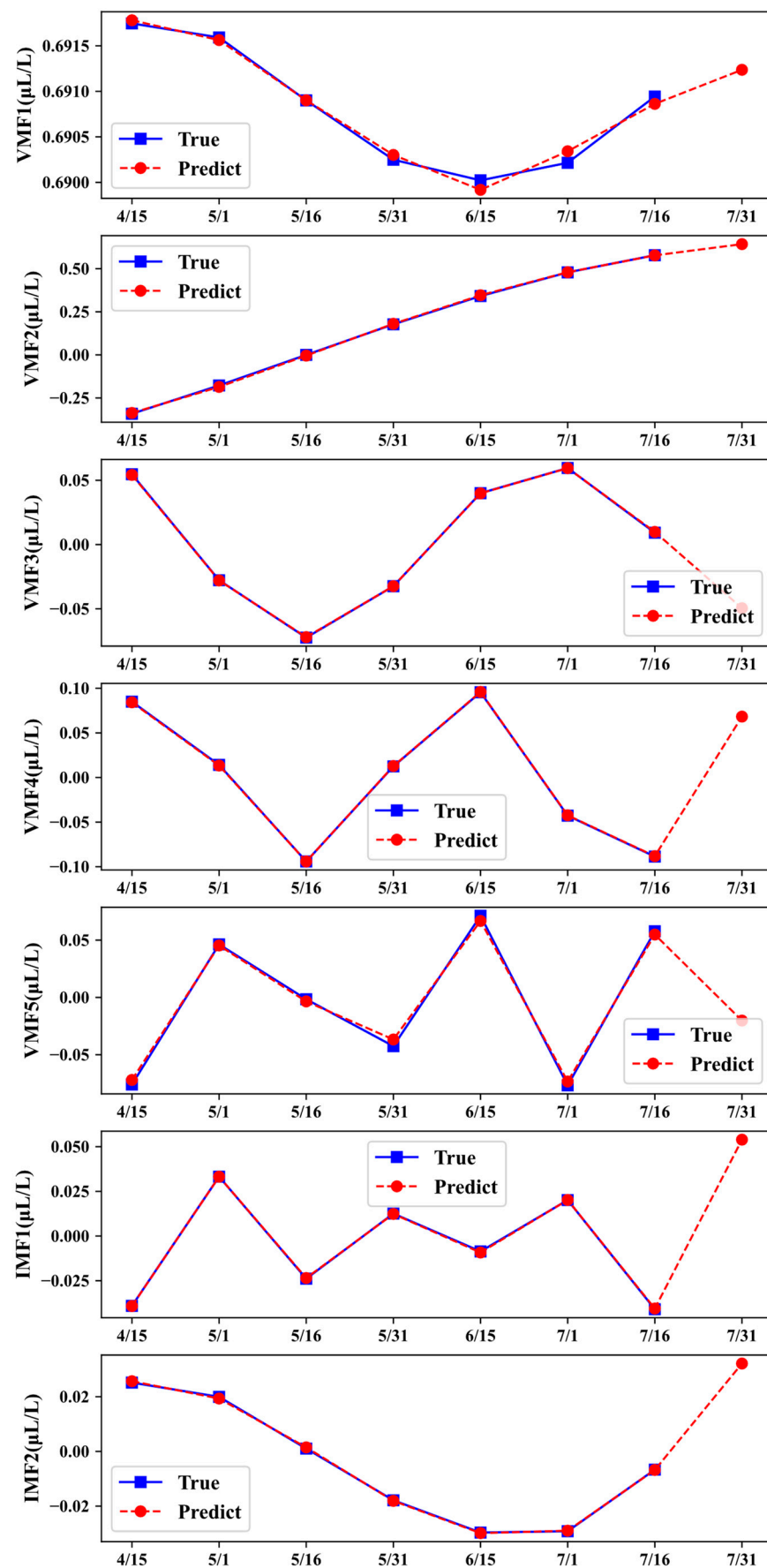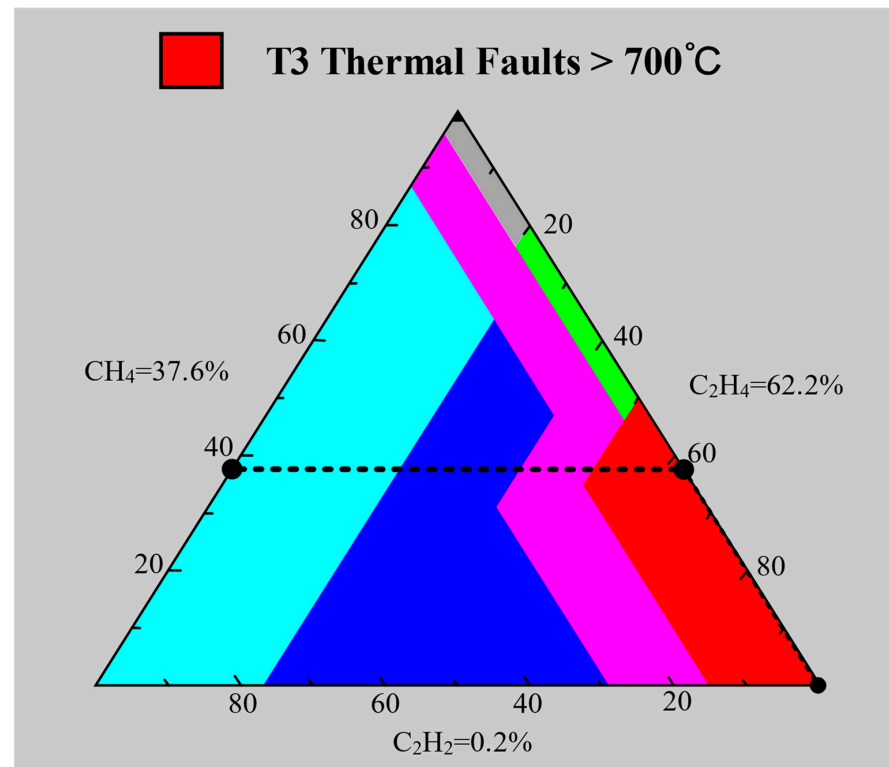
Figure 10. The prediction results for the acetylene sub-sequence on 31 July.

Table 4. Predicted and true values of characteristic gases in Case One (μL/L).

	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Predicted Value	71.3	224.6	153.0	372.3	1.3
True Value	71.2	224.9	152.6	371.4	1.4

Figure 11. Diagnostic result of Duval's triangle law in Case One.
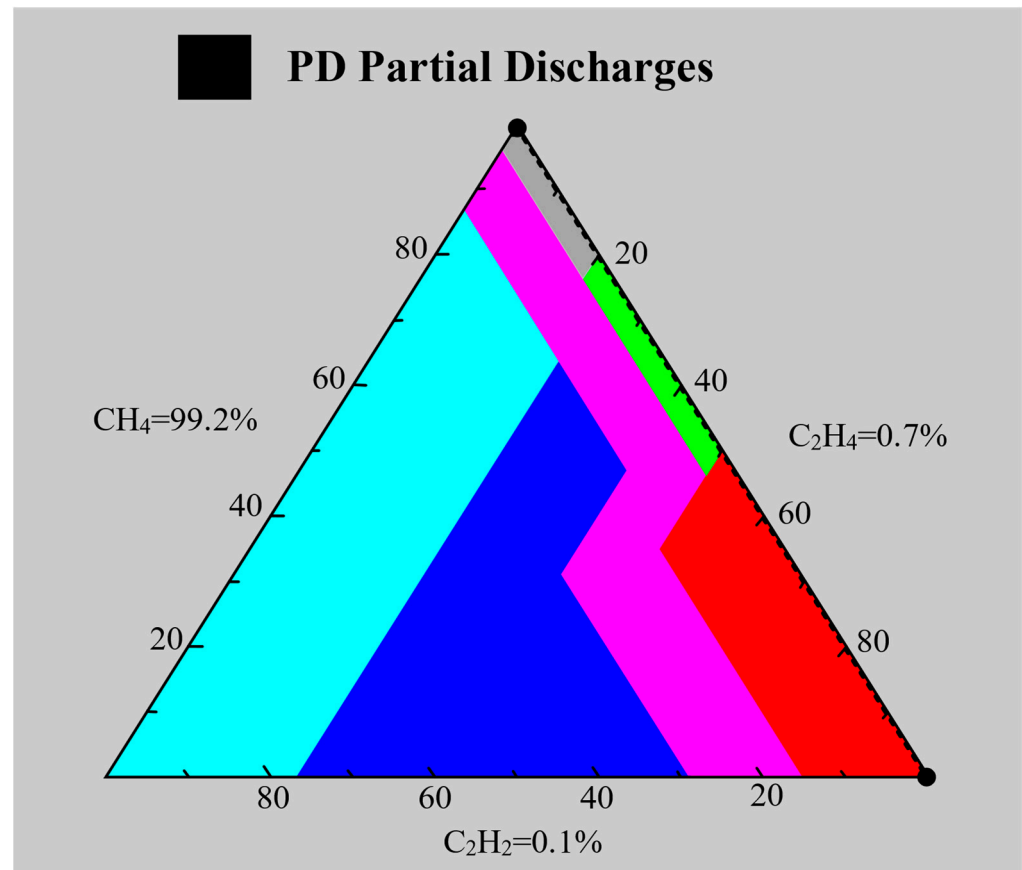
3.2. Case Two

In this case, the characteristic gases were predicted based on the data in the literature [29], and the operation process was consistent with the steps in Case One, which will not be described in detail. The final fault type of this transformer is partial discharge. Prior to this, sampling every 12 h from 4 May 2021 to 2 December 2021 and producing a total of 424 sets of sample points, due to the relative sufficiency of sample data, we take the last 30 sets of samples as a test set, and the rest for the training set. Based on the steps in Case One, the quadratic modal decomposition of the characteristic gases, the optimization of the prediction model, the prediction of the subsequences, and the superposition of the final subsequence values are carried out, respectively, and the prediction value of each characteristic gas on 2 December is obtained, whose results are shown in Table 5.

Table 5. December 2nd gas predicted and actual values in Case Two (μL/L).

	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Predicted Value	531.22	124.4	21.41	0.89	0
True Value	530.41	124.81	21.2	0.83	0

The predicted values of each characteristic gas in the Table 4 are imported into the ISSA-SVM fault diagnosis model, and the model predicts the fault result as partial discharge, which is consistent with the actual situation. The fault type judged using Duval's triangle law is located in the black area at the top of the triangle, which is a partial discharge fault,

and the diagnostic diagram is shown in Figure 12, which is consistent with the actual diagnostic results. The code obtained by utilizing the three-ratio method is 000, which has no matching fault when checking the fault type table.

Figure 12. Diagnostic result of Duval's triangle law in Case Two.

The analysis of the results in Case Two shows that there are limitations in the traditional fault diagnosis methods; for example, the three-ratio method struggles to determine the transformer fault types in this case, while the results of the state prediction model in this paper are consistent with the actual situation, which shows the excellent performance of the combined model.

3.3. Case Three

The data in Case Study 3 are sourced from the A-phase power transformer of Transformer #1 at a 500 kV substation in a certain power supply bureau of the Southern Power Grid in China [30]. An abnormal total hydrocarbon, exceeding the standard, began to appear on 22 January 2010. No issues were found during the preventive electrical tests but enhanced monitoring was implemented thereafter. The transformer experienced a fault on 23 July, and the gas concentration data on the day of the fault and prior to it are shown in Table 6. The diagnosed fault type was determined to be a medium-temperature overheating fault.

According to the same process in the previous case, the sequence decomposition, prediction, and sequence superposition steps were carried out for the characteristic gases, such as H_2 and CH_4, respectively, and the concentration values of each gas predicted by the SSA-BiLSTM model on 23 July are shown in Table 7.
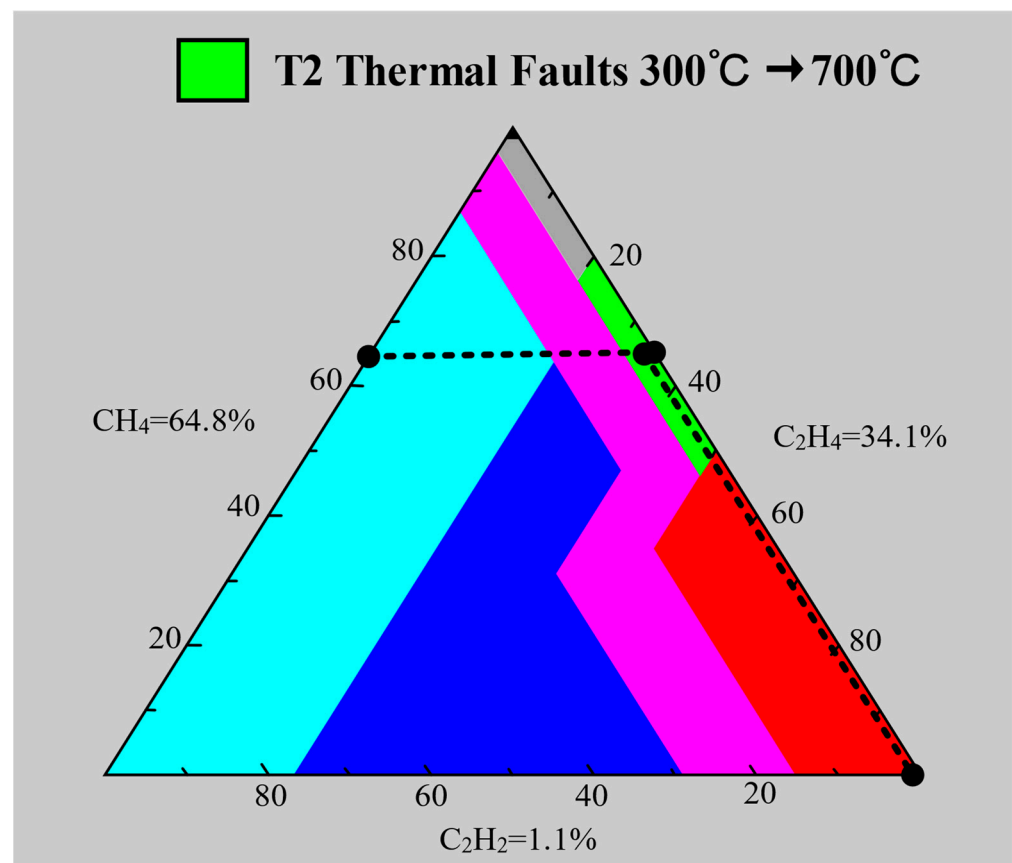
Table 6. Characteristic gas concentration values in Case Three (μL/L).

Date	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
6 May 2010	17.27	121.85	38.62	69.81	0.73
19 May 2010	17.88	137.42	40.37	74.72	0.82
31 May 2010	9.73	106.42	34.12	62.03	0.66
12 June 2010	21.43	140.95	39.38	75.65	1.7
24 June 2010	23.31	125.2	35.73	68.74	1.7
8 July 2010	10.01	100.68	35.1	64.53	1.46
23 July 2010	22.10	129.70	37.44	72.52	1.50

Table 7. The predicted and actual values of the characteristic gas on 23 July (μL/L).

	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Predicted Value	18.65	146.01	42.13	79.24	1.48
True Value	22.10	129.70	37.44	72.52	1.50

Taking the values of each gas concentration in Table 7 as inputs, the output of the ISSA-SVM fault diagnosis model in this paper diagnosed the fault type on 23 July as a medium-temperature overheating fault, which is the same as the actual fault situation on that day. The result graph judged via Duval's triangle law is shown in Figure 13, from which it can be seen that the fault point is located in the green area of the graph, which belongs to the medium-temperature fault. And after coding the data in Table 7 via the triple ratio method, the corresponding code is 021, and the diagnosed fault type can be determined to be the medium-temperature overheating fault.

Figure 13. Diagnostic result of Duval's triangle law in Case Three.

The results obtained from both the state prediction combined model and the traditional fault diagnosis methods are, in this case, the same as the actual fault types. Combining the fault prediction results of the three cases fully confirms that the combined model constructed in this paper has high accuracy in predicting the fault types of transformers.

4. Conclusions

Traditional gas prediction and fault diagnosis methods struggle to meet the demand of accurately determining the multiple fault types of oil-immersed transformers. Machine learning, as a tool that can quickly process a large amount of complex data and mine the relationships between the data, has been widely used in the field of gas prediction and the fault diagnosis of transformers. The transformer state prediction model constructed in this paper consists of a combination of two parts: gas prediction and fault diagnosis. The gas prediction model part adopts the optimal variational mode decomposition and CEEMDAN decomposition of residual sequences, which greatly reduces the noise contained in the original gas sequences, making it more accurate in predicting the future gas concentration and providing a good foundation for the fault diagnosis. In fault diagnosis, the improved sparrow search algorithm is first used to optimize the hyperparameters of the SVM so as to improve the accuracy of the classification of fault types; then, the future concentration values of each characteristic gas in the gas prediction model are imported, which can be used to classify and diagnose the possible future faults of the transformer. Based on the results of this paper, the main conclusions are as follows:

1. In Case One, the SSA-BiLSTM gas prediction model with quadratic modal decomposition is used to predict the concentration of each subsequence of methane and other gases. The results show that the model accurately predicts the concentration values on the day of several sampling dates before the fault occurs, and on the day of the fault, the predicted values of the concentration are very close to the actual values. The prediction results for Case Two are also very close to the real values. In Case Three, the deviation of the prediction results is larger than that for the other two cases due to the smaller amount of gas concentration data, but the overall prediction accuracy is very high, which demonstrates the excellent prediction performance of the gas prediction model proposed in this paper.

2. The improved sparrow search algorithm enhances the optimization ability and speed and is able to jump out of the local optimum quickly, making the SVM fault diagnosis algorithm's performance better. The predicted values of each characteristic gas in the model are imported into the ISSA-SVM fault diagnosis model, which is able to judge the possible future faults of the transformer. The results of the three cases prove that the combined transformer state prediction models constructed in this paper all accurately predict the results (i.e., they are consistent with the actual fault types), showing that the ISSA-SVM model has a high accuracy rate for transformer fault diagnosis.

3. Compared with machine learning fault diagnosis algorithms, traditional diagnostic methods such as Duval's triangle rule and the three-ratio method have certain limitations, which make it difficult to determine the actual types of faults occurring inside the transformer under the action of various factors, such as complex environment, accurately. It can be seen from the results of Case Two that sometimes, the fault type does not match the gas ratio after the encoding of the three-ratio method, so it is difficult for this method to be widely used in fault diagnosis. The oil-immersed transformer state prediction method based on the combined model of gas prediction and fault diagnosis proposed in this paper demonstrates good performance in judging the fault types for the three cases, and it will also have a positive impact on the operation and maintenance of transformers in the future.

References

1. Li, J.; Li, G.; Hai, C.; Guo, M. Transformer Fault Diagnosis Based on Multi-Class AdaBoost Algorithm. *IEEE Access* **2022**, *10*, 1522–1532. [CrossRef]
2. Jiang, J.; Chen, R.; Chen, M.; Wang, W.; Zhang, C. Dynamic Fault Prediction of Power Transformers Based on Hidden Markov Model of Dissolved Gases Analysis. *IEEE Trans. Power Deliv.* **2019**, *34*, 1393–1400. [CrossRef]
3. Ma, X.; Hu, H.; Shang, Y. A New Method for Transformer Fault Prediction Based on Multifeature Enhancement and Refined Long Short-Term Memory. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2512111. [CrossRef]
4. Wang, Y.; Zhang, P.; Qi, B.; Li, C.; Su, H.; Wang, Q. State Prediction of Power Transformer Based on Grey-Lagrange Method with Weighted Coefficient of Variation. In Proceedings of the 2018 Condition Monitoring and Diagnosis (CMD), Perth, WA, Australia, 23–26 September 2018; pp. 1–6.
5. Mishra, D.; Pradhan, A.K.; Baral, A.; Chakravorti, S. Reduction of time domain insulation response measurement duration for fast and effective diagnosis of power transformer. *IEEE Trans. Dielectr. Electr. Insul.* **2018**, *25*, 1932–1940. [CrossRef]
6. Benmahamed, Y.; Teguar, M.; Boubakeur, A. Application of SVM and KNN to Duval Pentagon 1 for transformer oil diagnosis. *IEEE Trans. Dielectr. Electr. Insul.* **2023**, *24*, 3443–3451. [CrossRef]
7. Dai, J.; Song, H.; Sheng, G.; Jiang, X. Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network. *IEEE Trans. Dielectr. Electr. Insul.* **2017**, *24*, 2828–2835. [CrossRef]
8. Guo, J.; Xie, Z.; Qin, Y.; Jia, L.; Wang, Y. Short-Term Abnormal Passenger Flow Prediction Based on the Fusion of SVR and LSTM. *IEEE Access* **2019**, *7*, 42946–42955. [CrossRef]
9. Guo, C.; Dong, M.; Wu, Z. Fault Diagnosis of Power Transformers Based on Comprehensive Machine Learning of Dissolved Gas Analysis. In Proceedings of the IEEE 20th International Conference on Dielectric Liquids (ICDL), Roma, Italy, 23–27 June 2019; pp. 1–4.
10. Qi, B.; Wang, Y.; Zhang, P.; Li, C.; Wang, H. A novel deep recurrent belief network model for trend prediction of transformer DGA data. *IEEE Access* **2019**, *7*, 80069–80078. [CrossRef]
11. Han, X.; Huang, S.; Zhang, X.; Zhu, Y.; An, G.; Du, Z. A transformer condition recognition method based on dissolved gas analysis features selection and multiple models fusion. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106518. [CrossRef]
12. Ali, M.S.; Abu Bakar, A.H.; Omar, A.; Abdul Jaafar, A.S.; Mohamed, S.H. Conventional methods of dissolved gas analysis using oil-immerse-d power transformer for fault diagnosis: A review. *Electr. Power Syst. Res.* **2023**, *216*, 109064. [CrossRef]
13. Code, P.; Prix, C. *Mineral Oil-Impregnated Electrical Equipment in Service–Guide to the Interpretation of Dissolved and Free Gases Analysis*; IEC Publication 60599; British Standards Institution: London, UK, 2007.
14. Duval, M. A review of faults detectable by gas-in-oil analysis in transformers. *IEEE Electr. Insul. Mag.* **2021**, *18*, 8–17. [CrossRef]
15. Rogers, R.R. IEEE and IEC Codes to Interpret Incipient Faults in Transformers, Using Gas in Oil Analysis. *IEEE Electr. Insul. Mag.* **1978**, *EI-13*, 349–354. [CrossRef]
16. Xing, Z.; He, Y.; Wang, X.; Shao, K.; Duan, J. VMD-IARIMA-Based Time-Series Forecasting Model and its Application in Dissolved Gas Analysis. *IEEE Trans. Dielectr. Electr. Insul.* **2023**, *30*, 802–811. [CrossRef]
17. Zhang, W.; Liu, J.; Jiang, J.; Zhang, X.; Fan, L. Prediction of concentration for dissolved gas in oil based on CEEMDAN and TCN. *Electr. Power Eng. Technol.* **2024**, *30*, 192–200.
18. Huang, X.; Wang, X.; Tian, Y. Research on Transformer Fault Diagnosis Method based on GWO Optimized Hybrid Kernel Extreme Learning Machine. In Proceedings of the 2018 Condition Monitoring and Diagnosis (CMD), Perth, WA, Australia, 23–26 September 2018; pp. 1–5.
19. Das, S.; Paramane, A.; Chatterjee, S.; Rao, U.M. Accurate Identification of Transformer Faults From Dissolved Gas Data Using Recursive Feature Elimination Method. *IEEE Trans. Dielectr. Electr. Insul.* **2023**, *30*, 466–473. [CrossRef]
20. Das, S.; Paramane, A.; Chatterjee, S.; Rao, U.M. Sensing Incipient Faults in Power Transformers Using Bi-Directional Long Short-Term Memory Network. *IEEE Sens. Lett.* **2023**, *7*, 7000304. [CrossRef]
21. Baiju, S.S.; Adarsh, S. Root Cause Analysis in Power Transformer Failure with Improved Intelligent Methods. In Proceedings of the 2022 IEEE International Power and Renewable Energy Conference (IPRECON), Kollam, India, 16–18 December 2022; pp. 1–6.
22. Mounir, N.; Ouadi, H.; Jrhilifa, I. Short-term electric load forecasting using an EMD-BI-LSTM approach for smart grid energy management system. *Energy Build.* **2023**, *288*, 113022. [CrossRef]

23. Qu, L.; Zhou, H. The Multi-class SVM Is Applied in Transformer Fault Diagnosis. In Proceedings of the 2015 14th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), Guiyang, China, 18–24 August 2015; pp. 477–480.
24. Li, J.; Lei, Y.; Yang, S. Mid-long term load forecasting model based on support vector machine optimized by improved sparrow search algorithm. *Energy Rep.* **2022**, *8*, 491–497. [CrossRef]
25. Gu, Q.; Tian, J.; Li, X.; Jiang, S. A novel Random Forest integrated model for imbalanced data classification problem. *Knowl.-Based Syst.* **2021**, *250*, 109050. [CrossRef]
26. Zhang, H.; Zhang, Y. An Improved Sparrow Search Algorithm for Optimizing Support Vector Machines. *IEEE Access* **2023**, *11*, 8199–8206. [CrossRef]
27. Wu, C.; Lin, J.; Yu, Z.; Yang, J.; Liu, X. Thermal State Prediction of Transformers Based on ISSA-LSTM. In Proceedings of the 2022 5th International Conference on Energy, Electrical and Power Engineering (CEEPE), Chongqing, China, 22–24 April 2022; pp. 160–165.
28. Ding, C.; Ding, Q.C.; Wang, Z.Y.; Zhou, Y.Y. Fault diagnosis of oil-immersed transformers based on the improved sparrow search algorithm optimised support vector machine. *IET Electr. Power Appl.* **2021**, *16*, 985–995. [CrossRef]
29. Ding, C.; Ding, Q.C.; Feng, L.; Wang, Z.L. Prediction Model of Dissolved Gas in Transformer Oil Based on VMD-SMA-LSSVM. *IEEJ Trans. Electr. Electron. Eng.* **2022**, *17*, 1432–1440. [CrossRef]
30. SiMa, L. Research on Fault Diagnosis and Prediction Method of Power Transformer Based on Improved Support Vector Machine. Ph.D. Thesis, Department of Electrical Engineering, Wuhan University, Wuhan, China, 2012.