Pre-processing approach for de-noising on-line oil chromatography data based on self-adapting wavelet analysis

Chunyan Gong¹, Wei Hu¹, Xiaohong Wu¹, Bing Lu¹, Yiming Zheng²

¹State Grid Hangzhou Power Supply Company, Hangzhou, People’s Republic of China
²State Grid Zhejiang Electric Power Research Institute, Hangzhou, People’s Republic of China

E-mail: Yiming_zheng@zj.sgcc.com.cn

Abstract: Due to the influence of the external environment and the performance of measuring equipment, the on-line oil chromatography data contains obvious noise which makes the signal to oscillate. The monitoring data is difficult to be directly applied to the equipment state analysis. A novel wavelet-based de-noising method is proposed for pre-processing the on-line oil chromatography data. By analysing the characteristics of on-line oil chromatography data, the method of determining the decomposition layer number based on the probability distribution of wavelet coefficients and the method of determining the threshold value based on outliers conservation are proposed. The improved wavelet de-noising method is applied to analysing the on-line oil chromatography data of a defective ultra-high voltage (UHV) reactor. The results show that the proposed method is feasible and effective.

1 Introduction

As the pillar of the national economy, the power industry is in continuous progress and innovation. Safe and stable operation of the power system is determined by the reliability of power equipment. After synchronisation of Sanhua power networks, progressively more electric power devices were put into operation. Thus, operations and maintenance of power equipment become increasingly significant [1, 2]. Detections of dissolved gas are the main measure for fault diagnosis and life prediction in large oil-immersed power equipment [3]. At the present time, the off-line dissolved gas analysis is the main method to evaluate the oil-immersed power equipment states [4]. However, a sampling interval of off-line detection is too large to pay close attention to the oil chromatogram changing and detect the latent defects in oil-immersed power equipment. On-line oil chromatogram monitoring makes up these defects, and can realise real-time insulation monitoring, providing a wealth of information for fault diagnosis.

With the continuous innovation of oil chromatography on-line monitoring technology, ever more oil-immersed equipment were equipped with on-line oil chromatogram monitoring device [5, 6].

Nevertheless, due to the influence of the external environment and the performance of measuring equipment, obvious errors and signal oscillation can be observed in on-line monitoring data. A common approach to reducing the influence is to average the monitoring data of each day. Such an approach can effectively diminish the noise and oscillation. However, the monitoring data number will decrease to one per day, and the data continuity of the adjacent 2 days will also be man-made destructed. Moreover, the advantage of intensive sampling points is also impaired.

This paper puts forward a wavelet de-noising method for pre-processing on-line oil chromatography data. This method was applied to a defect UHV reactor, and compared with traditional methods. Theoretical analysis and practical application show that the proposed self-adapting wavelet de-noising method can effectively improve the quality of on-line monitoring data, and the de-noised data can be used for fault diagnosis.

2 On-line oil chromatogram

On-line oil chromatography is a kind of oil monitoring means. Sampling density is determined according to the voltage level and the importance degree. For ultra-high-voltage equipment, sample interval is always set as 15 min. On-line oil chromatography data ($C_2H_2$) of an ultra-high-voltage reactor is drawn in Fig. 1.

As can be seen from Fig. 1, the acetylene content in this reactor is abnormal. After deleting the invalid data, daily sampling points is about 80–90, which can reflect the real-time insulation condition. Owing to the environmental and the performance of measuring equipment, the data oscillates obviously. In addition, the gas circulation also has some impact on the measurement results. Theoretically, errors caused by the environment and the measurement follow a normal distribution. Since the insulating oil circulates in oil-immersed equipment, the error caused by oil circulation can be regarded as normal distribution as well. Errors caused by different factors are independent of each other; thus, mathematical methods can be used to eliminate the noise after understanding the characteristics of the data.

3 Self-adapting wavelet de-noising method

3.1 Decomposition layer number

The decomposition layer number affects the quality of the de-noised signal. The greater the layer number is, the more obvious the difference in the characteristics of noise and signal perform. On the other hand, more decomposition layers lead to a larger signal loss in thresholding, making the reconstruction error greater.

Fig. 1 Typical on-line oil chromatography data
For a set of \( N \) data signal, the theoretical maximum decomposition layer number \( K \) can be calculated by

\[
K \leq \log_2 N
\]  

(1)

In order to retain the necessary detail information, the decomposition layer number should be limited. Since the noise follows a normal distribution, a method for selecting the decomposition layers number is presented based on the probability distribution of the wavelet coefficients. The flowchart is shown in Fig. 2.

In Fig. 2, wavelet coefficients of each layer should go through the normal distribution test. If the coefficients follow a normal distribution, then continue to next decomposition, otherwise stop the decomposition. On the basis of this process, it can be ensured that the eliminated signal follow a normal distribution.

This method can ensure adequate decomposition layers, and retain the effective signal as much as possible. Since high order \( d_k \) will contain a part of the effective information, the decomposition layer number obtained by this method will not exceed the maximum decomposition layer number \( K \). Kolmogorov–Smirnov (K–S) test can be chosen as the normal distribution test method.

3.2 Thresholds determination method

After determining the decomposition layer number, it can be ensured that the wavelet coefficients follow a normal distribution; however, there are still singular values. These singular values correspond to valid signals. A self-adapting threshold processing method is proposed here to retain the valid information, which can eliminate the noise as far as possible. In order to achieve an improved de-noising effect, the hypothesis testing requirements can be relaxed in the process of determining the layer number.

The threshold selection process is shown in Fig. 3.

The proposed threshold selecting method is a self-adapting method based on data characteristics. It can ensure that the singular value (valid information) in wavelet coefficients is eliminated.

4 Case study

4.1 On-line monitoring of oil chromatogram de-noising

Take the on-line monitoring of acetylene content data in Fig. 1 as an example for de-noising. Determine the decomposition layer number by the process specified in Fig. 2. The K–S testing results of wavelet coefficients of 1–8 layers are shown in Table 1. The confidence level is set to 0.01. Wavelet coefficients histograms are shown in Fig. 4.

Wavelet coefficients of 1–7 layers can go through normal distribution test, while the wavelet coefficients of the eighth layer cannot. From the histogram, it can be seen clearly that \( d_8 \) shows

Table 1 K–S test results for each wavelet coefficient (confidence level \( \alpha = 0.01 \))

| Wavelet coefficients | \( p \) value | \( h \) |
|----------------------|--------------|------|
| \( d_1 \)            | 0.0806       | 0    |
| \( d_2 \)            | 0.0854       | 0    |
| \( d_3 \)            | 0.0485       | 0    |
| \( d_4 \)            | 0.628        | 0    |
| \( d_5 \)            | 0.578        | 0    |
| \( d_6 \)            | 0.4046       | 0    |
| \( d_7 \)            | 0.6633       | 0    |
| \( d_8 \)            | 0.0005       | 1    |
the obvious deviation compared to normal distribution. Therefore, the decomposition layer number is determined as 7.

Then determine the threshold values using the above proposed method. The calculated threshold values are shown in Table 2.

Process all the wavelet coefficients by the threshold values in Table 2. By reconstructing the signal, the de-noised on-line oil chromatography data can be obtained.

4.2 De-noising effect evaluation

The de-noised data and the original data are drawn in Fig. 5. Oscillations can no longer be observed in a de-noised curve while the trends of the original data are conserved. Furthermore, the filtered signal is drawn in Fig. 6 along with the histogram.

The filtered signal follows a normal distribution and the \( p \) value of the K–S test is 0.7444.

As a comparison, the decomposition layer number is calculated according to (1). The data length is 5964, thus the decomposition layer number is set to 12. The heursure method is selected as the threshold determining method. Similarly, the filtered signal and its histogram are plotted in Fig. 7.

It can be seen that the filtered signal deviates slightly from a normal distribution, especially from 6th to 27th September. Outlier values in the histogram also increased significantly. A \( p \) value of the K–S test is 0.1716, which can also pass the test. However, its value is less than the self-adapting method. It can be proved that the proposed wavelet de-noising method can improve the de-noise quality of on-line monitoring oil chromatography data.

4.3 Comparison between the proposed and traditional methods

Commonly, the processing method for on-line oil chromatography data is day-average. The de-noised and averaged on-line oil chromatography data are shown in Fig. 7 for further comparison (Fig. 8).

It is obvious that these two curves are in good agreement. A certain degree of oscillation can be observed in the day-average curve and, of course, the data quantity is less. The de-noised curve is relatively smooth and data is more intensive, which can reflect more details in oil chromatography data.

The gas content dissolved in the oil is a cumulative amount. However, the growth rate can reflect the oil deterioration more clearly. Based on the de-noised and the day-average oil chromatography data, the first-order differences are calculated. The results are shown in Fig. 9.

It can be concluded that:

i. The wavelet de-noising method can warn earlier than the day-average method when the acetylene content increases.

ii. Uprushes are detected in the first-order differences of day-average data on 30th September, 13th October and 2nd November; however, no significant abnormality is found in the

| Wavelet coefficients | Threshold |
|----------------------|-----------|
| \( d_1 \)            | 0.1177    |
| \( d_2 \)            | 0.1837    |
| \( d_3 \)            | 0.3272    |
| \( d_4 \)            | 0.4261    |
| \( d_5 \)            | 0.3625    |
| \( d_6 \)            | 0.2877    |
| \( d_7 \)            | 0.2797    |

Table 2 Threshold value of the wavelet coefficients based on the outlier conservation

Fig. 5 Comparison between the original and the de-noising on-line oil gas chromatography

Fig. 6 Noise signal and its histogram (proposed method)

Fig. 7 Noise signal and its histogram (decomposition level = 12, heursure)

Fig. 8 Comparison between the wavelet de-noising method and the conventional method
original acetylene content curve. These uprushes are caused by errors.

iii. Wavelet de-noising method can use the data more effectively than the day-average method, the time of sudden increase in acetylene content is also more accurate.

It can be concluded that compared to the day-average method, the wavelet de-noise method can position the time of uprushes more timely and accurately.

5 Conclusions

i. The decomposition layer number determining method based on the probability distribution of wavelet coefficients test the wavelet coefficients of each layer. Thus, it can be ensured that the noise is eliminated and effective information is retained.

ii. The threshold determination method based on outlier conservation can detect singular values in wavelet coefficients, which represent the effective information. When selecting the threshold, these values are retained to ensure the effectiveness of the de-noised data.

iii. The self-adapting wavelet de-noising method proposed in this paper can effectively eliminate the oscillation phenomenon caused by the noise and error in the original on-line monitoring data, and can maintain the overall trend of the original data.

iv. The eliminate signal by the proposed self-adapting wavelet de-noising method follows normal distribution better than the traditional wavelet method, and the de-noising effect is better.

v. The first-order difference of de-noise on-line monitoring data can reflect the changes of dissolved gas in oil better, and can indicate the mutation more timely and more accurately. The proposed method can be applied to fault diagnosis and life prediction, and has important practical significance.

6 References

[1] Sun, C.: ‘Present situation and development of condition on-line monitoring and diagnostic technology for power transmission and transformation equipment’, Electr. Power, 2005, 38, (2), pp. 1–7

[2] Liu, L., He, W., Liu, Y., et al.: ‘Analysis software design of on-line monitoring and fault diagnosis system for power transmission equipment’, Comput. Syst. Appl., 2011, 8, (8), pp. 27–32

[3] Liang, Y., Cao, Y.: ‘Application of transformer oil chromatographic analysis technology to equipment condition maintenance’, Transformer, 2011, 12, (12), pp. 67–71

[4] Guan, G., He, J.: ‘On-line insulation monitoring techniques and condition maintenance of electric equipment’, Transformer, 2004, 39, (21), pp. 46–50

[5] Zhang, H., Zhu, S., Zhang, Y., et al.: ‘Research and implementation of condition-based maintenance technology system for power transmission and distribution equipment’, Power Syst. Technol., 2009, 33, (1), pp. 70–73

[6] Guo, B., Yang, X.: ‘Analysis on development and application of monitoring technique for electric power devices in chin’, Power Syst. Technol., 2004, 23, (8), pp. 65–68

Fig. 9 First-order differences of oil chromatography data