Journal of Energy, Environmental & Chemical Engineering
2018; 3(4): 54-60
http://www.sciencepublishinggroup.com/j/jeece
doi: 10.11648/j.jeece.20180304.11
ISSN: 2637-4331 (Print); ISSN: 2637-434X (Online)

Sulfur Dioxide Detection Signal Denoising Based on Support Vector Machine

Zhifang Wang, Shutao Wang*

Institute of Electrical Engineering, Yanshan University, Qinhuangdao, China

Email address:

wangzhifang0119@163.com (Zhifang Wang), wangshutao@ysu.edu.cn (Shutao Wang)

*Corresponding author

To cite this article:

Zhifang Wang, Shutao Wang. Sulfur Dioxide Detection Signal Denoising Based on Support Vector Machine. *Journal of Energy, Environmental & Chemical Engineering.* Vol. 3, No. 4, 2018, pp. 54-60. doi: 10.11648/j.jeece.20180304.11

Received: January 22, 2019; **Accepted**: February 26, 2019; **Published**: March 19, 2019

Abstract: A system for detecting sulfur dioxide (SO_2) based on differential optical absorption spectrometry theory was studied. The detection system can eliminate the noise from light source and light path by using the double optical path. Background noise was generated by the photoelectric device. It also effects the quantitative analysis. The Support Vector Machine (SVM) is proposed to process the SO_2 ultraviolet absorption spectrum. The SO_2 ultraviolet absorption spectra at 220nm-340nm were obtained by using the SO_2 detection system in this article. Then the spectral was denoised by the SVM. The experimental results showed that the absorption line was more smoothness after denoising by the SVM, and the SNR and mean square error were 48.9398 and 1×10^{-7}, respectively. The de-noising data was applied to the SO_2 detection system, the linearity of the measurement was good with the coefficients of more than 0.9971. Compare the result with the wavelet and Empirical Mode Decomposition (EMD) denoising methods, which illustrates that SVM has better effects. It shows that the SVM method applied to noise reduction of SO_2 detection system is superior.

Keywords: Sulfur Dioxide, Denoising, Support Vector Machine, Wavelet, Empirical Mode Decomposition

1. Introduction

With developing of industry, air pollution becomes more and more serious. SO_2 is one of the main pollutants in the atmosphere. The study found that increasing SO_2 levels caused respiratory and cardiovascular diseases [1, 2], exacerbate the formation of acid rain [3], and harm the growth of crops. There are many detection methods to detecting the SO_2[4], such as conductometry, chemiluminescence and spectroscopy. Spectroscopic method has the advantage of high measurement accuracy, large measurement range, simple operation, and low cost, so it is widely used to detect the SO_2 for these advantages.

The concentration of trace gases in the atmosphere is much lower than 1% [5]. SO_2 is a trace gas, so the signal of SO_2 is weak, and it is easily affected by various noises of the instrument and unavoidable environmental factors [6]. In recent years, many denoising methods have emerged. Li uses wavelet method to denoise SO_2 spectrum [7], but wavelet denoising method is not effective when dealing with nonlinear signals [8, 9]. Li uses the EMD method for SO_2 fluorescence spectral denoising [10], EMD denoising method can adaptively decompose the signal into intrinsic mode functions (IMF). After the signal is decomposed, SO_2 fluorescence spectral signal denoising is implemented by removing high frequency noise signals and reconstruct [11]. But it removes useful information in high frequency signals and ignores noise information in low frequency signals. Wang research used three algorithms of wavelet filtering, EMD filter and Boxcar filter to extract and recover the SO_2 fluorescence signal drowned in the noise floor. The high-precision measurements are needed with the development of industry, and it puts forward higher request for denoising technique. Support Vector Machine (SVM) is an algorithm based on statistical theory and probability theory. SVM is used as a classifier in some studies [12-18]. SVM is also used for signal denoising based on strong generalization ability and global optimization. [19-22]. Sidheswar designed and developed an efficient hybrid image denoising method based on SVM, based on a predefined threshold, SVM is used to classify the patches into two classes such as texture patches and flat patches [19]. Hamid Reza Shahdoosti used the wavelet algorithm to extract the features of the unmixed wavelet

segments and used the twin support vector machine to achieve image denoising [20]. Chen proposed a method for signal filtering based on SVM, the effects of parameter on kernel function are analyzed in time domain and frequency domain respectively [21]. Ji proposed a hardware implementation method of FIR filter based on improved SVM design method [22]. SVM is used for classifiers or isn't used for SO_2 denoising. In this paper, SVM is used for SO_2 absorption spectrum denoising. A model is built by repeatedly training a part of the signal, and the remaining signals are used to test. The training model needs a small number of training samples to effectively remove the noise signal while preserving useful information.

2. Methodology and Detection System

In this section, SVM algorism principle will be introduced first. Then the SO_2 detection system is introduced. The performances of the detection system and its practical results are also presented.

2.1. SVM

The SVM was proposed by Vapnik in 1995[23]. SVM is a way of machine learning, it has good effect for small samples, nonlinear, high dimensional, etc. SVM is developed from solving linear problems, it can construct optimal hyperplane under the condition of linear and divisible. However, in practical application, most problems are nonlinear. Therefore, the nonlinearity of the input data map to a high-dimensional feature space. For example, given a set of array lengths n, which belong to $\mathbf{R}^d$:

$$(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), ..., (\boldsymbol{x}_n, y_n)$$

$$\boldsymbol{x}_i \in \mathbf{R}^d, y_i \in \{-1, +1\}, i = 1, ..., n \tag{1}$$

Where $x_i \in \mathrm{R}^d$ are input training samples, y_i output are training samples, d is dimension, i is the ith number in the array.

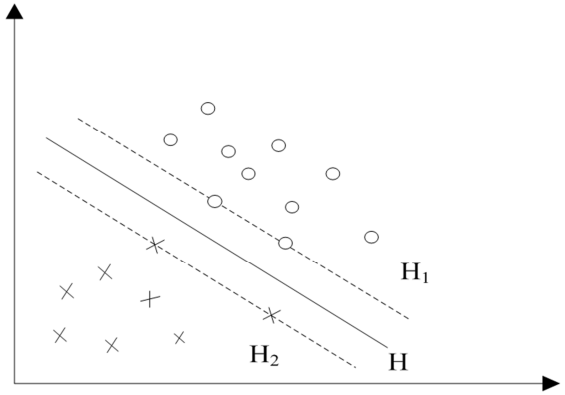


Figure 1. *Optimization classify hyperplane under linear condition.*

If the samples are separable, there is a classification hyperplane that separates the two types of samples. Cross and open circles represent two types of samples in Figure 1. The nearest point to the classified hyperplane is called the support vector. H is a classification hyperplane. Hyperplanes H_1 and H_2 are over two types of support vectors and are parallel to H, respectively. The distance between H_1 and H is equal to the distance between H_2 and H. The distance between H_1 and H_2 is called the classification interval.

If a hyperplane divides the data into two categories, divided as follows:

$$\begin{cases} (\boldsymbol{w} \bullet \boldsymbol{x}_i) + \mathrm{b} \geq 0, & y_i = +1 \\ (\boldsymbol{w} \bullet \boldsymbol{x}_i) + \mathrm{b} \leq 0, & y_i = -1 \end{cases} \tag{2}$$

where $\boldsymbol{w}$ is vector of hyperplane. $\boldsymbol{x}$ is input vector of training set. b is constant term of hyperplane. Hyperplane over two types of sample support vectors is defined as:

$$\begin{cases} (\boldsymbol{w} \bullet \boldsymbol{x}_i) + \mathrm{b} = +1 \\ (\boldsymbol{w} \bullet \boldsymbol{x}_i) + \mathrm{b} = -1 \end{cases} \tag{3}$$

The interval d between hyperplanes $\mathbf{H}_1$ and $\mathbf{H}_2$ can be obtained from Eq. (3):

$$d = \frac{2}{|\boldsymbol{w}|} \tag{4}$$

The regression function of classification hyperplane is defined as:

$$f(\boldsymbol{x}) = (\boldsymbol{w} \bullet \boldsymbol{x}) + \mathrm{b}, \boldsymbol{w} \in \mathbf{R}^d, \mathrm{b} \in \mathbf{R} \tag{5}$$

The optimal hyperplane has maximum margin between two classes. The optimal hyperplane problem is transformed into solving the quadratic optimization, and the slack variable is introduced [24]. The quadratic form can be represented as:

$$\min \frac{1}{2}\|\boldsymbol{w}\|^2 + \mathrm{C}\sum_{i=1}^{n} \xi_i + \xi_i^* \Bigg\}$$
$$s.t. \begin{cases} y_i - (\boldsymbol{w} \bullet \boldsymbol{x}_i) - \mathrm{b} \leq \varepsilon + \xi_i \\ (\boldsymbol{w} \bullet \boldsymbol{x}_i) + \mathrm{b} - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \tag{6}$$

where ξ_i, ξ_i^* are relaxation factors, C is the penalty factor, ε is the insensitivity coefficient, and *s.t.* is constraint.

For the complexity of calculations of the quadratic optimization, the Eq. (6) is transformed into a dual problem with Lagrange duality theory, and the Eq. (6) can be transformed as Eq. (7):

$$\begin{aligned} \mathrm{L}(\boldsymbol{w}, \xi_i, \xi_i^*, \alpha, \alpha^*, \mathrm{C}, \beta, \beta^*) = {} & \frac{1}{2}\|\boldsymbol{w}\|^2 + \mathrm{C}\sum_{i=1}^{n} \xi_i + \xi_i^* \\ & - \sum_{i=1}^{n} \alpha_i [(\boldsymbol{w} \bullet \boldsymbol{x}_i) + \mathrm{b} - y_i + \varepsilon + \xi_i] \\ & - \sum_{i=1}^{n} \alpha_i^* [y_i - (\boldsymbol{w} \bullet \boldsymbol{x}_i) - \mathrm{b} + \varepsilon + \xi_i^*] \\ & - \sum_{i=1}^{n} (\beta_i \xi_i + \beta_i^* \xi_i^*) \end{aligned} \tag{7}$$

Where α_i, α_i^*, β_i, β_i^* are all Lagrange multipliers. Then the Eq. (6) can be expressed as:

$$\begin{cases} \max\limits_{\alpha_i,\alpha_i^*} -\varepsilon \sum\limits_{i=1}^{l}(\alpha_i+\alpha_i^*)+\sum\limits_{i=1}^{l} y_i(\alpha_i-\alpha_i^*) \\ -\frac{1}{2}\sum\limits_{i,j=1}^{l}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)(\boldsymbol{x}_i\bullet\boldsymbol{x}_j) \\ s.t.\begin{cases} \sum\limits_{i=1}^{l}(\alpha_i-\alpha_i^*)=0 \\ \alpha_i,\alpha_i^*\in[0,C] \end{cases} \end{cases} \tag{8}$$

The expression of $f(\boldsymbol{x})$ in Eq. (5) is expressed as:

$$f(\boldsymbol{x})=\sum_{i=1}^{l}\beta_i(\boldsymbol{x}_i\bullet\boldsymbol{x})+b \tag{9}$$

β_i is non-support-vector. When the data set is certain, β_i=0. The $f(\boldsymbol{x})$ is represented by the remaining support vectors as:

$$f(\boldsymbol{x})=\sum_{i\in N}\beta_i(\boldsymbol{x}_i\bullet\boldsymbol{x})+b \tag{10}$$

Where N is a subset of the input data set. For a particular problem, a model for this problem can be determined by a subset of given data.

When the problem is nonlinear, Eq. (10) cannot accurately represent $f(\boldsymbol{x})$. The nonlinearity of the input data map to a high-dimensional feature space by using nonlinear mapping $\boldsymbol{\Phi}(\mathbf{x})$. In order to reduce the amount of calculation, the inner product operation of the high-dimensional feature space is converted into a function transport of the input space by using the kernel function K ($\boldsymbol{x}_i$, $\boldsymbol{x}$).

$$K(\boldsymbol{x}_i,\boldsymbol{x}_j)=(\Phi(\boldsymbol{x}_i)\bullet\Phi(\boldsymbol{x}_j)) \tag{11}$$

The expression of $f(\boldsymbol{x})$ of Eq. (10) can be expressed as:

$$f(\boldsymbol{x})=\sum_{i\in N}\beta_i K(\boldsymbol{x}_i,\boldsymbol{x})+b \tag{12}$$

The common kernel functions are: linear kernel function, polynomial kernel function, sigmoid kernel function and gaussian radical kernel function. The gaussian radical kernel function is better for problems with less a priori information [25]. The expression of the gaussian radical kernel function is expressed as:

$$K(\boldsymbol{x}_i,\boldsymbol{x}_j)=\exp\left(-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{2r^2}\right) \tag{13}$$

The basic concept of the SVM denoising method is to construct a good model structure. Constructed model is able to predict unknown data and remove noise. Simple models cannot accurately predict unknown data, and overly complex models contain interference signals. The width of the gaussian radical kernel function is determined by the parameter r. According to adjust the parameter r, the complexity of the model is controlled by using the SVM. The model can accurately describe characterizes of unknown data, and removes noise information.

2.2. SO_2 Detection System

SO_2 has three absorption peaks in the ultraviolet band, the first band: 320~390nm, SO_2 has a weak absorption; the second band: 240~320nm, SO_2 has a stronger absorption; the third band: 190~240nm, SO_2 has the strongest absorption.

The SO_2 detection system with differential optical absorption spectrometry was used to detect the SO_2 concentration. The system is mainly composed of light source, gas chamber, signal collect and processing part. The deuterium lamp (LHD30, Zolix) is used as excitation source, it has peak in 200~300nm and good continuity in 200~400nm. There are two chambers to form a differential detection. The SO_2 detection system is shown in Figure 2.

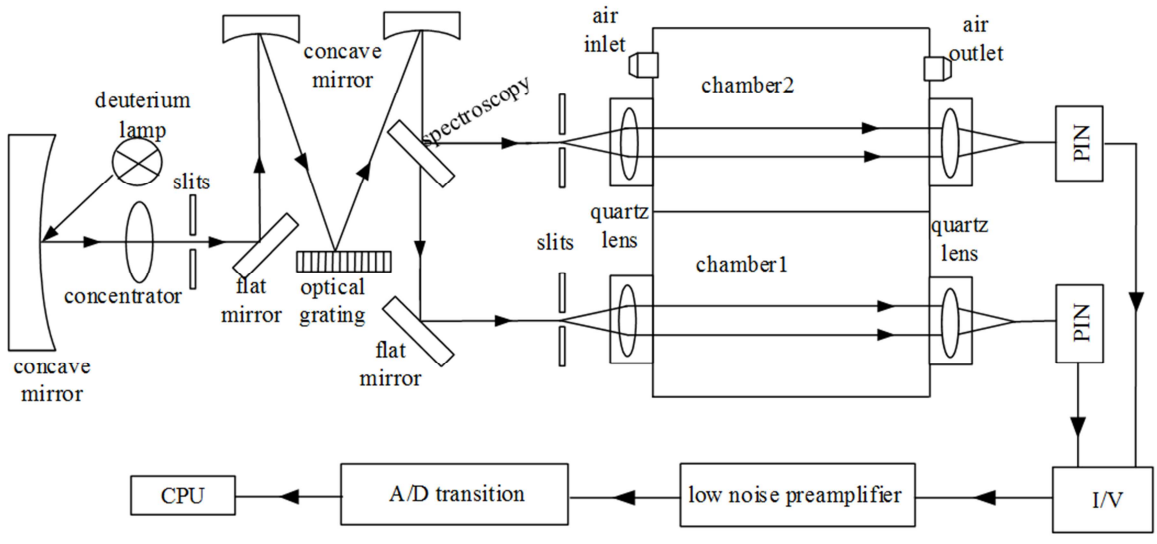


Figure 2. *SO_2 detection system.*

The ultraviolet absorption spectrum of SO_2 in the range of 220 nm to 340 nm is shown in Figure 3. It can be seen that there is noise in the ultraviolet absorption spectrum of SO_2, so SVM is used for denoising.

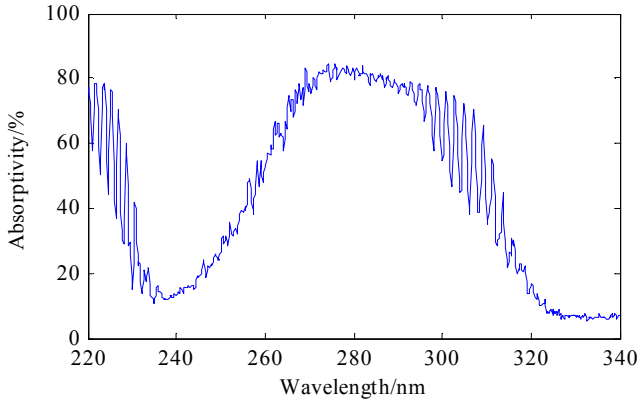


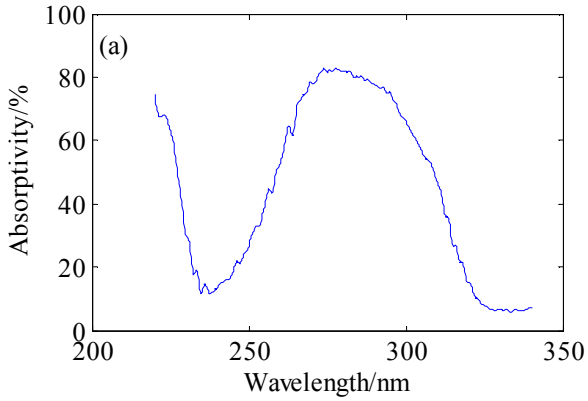
Figure 3. *The ultraviolet absorption spectrum of SO_2.*

3. Results

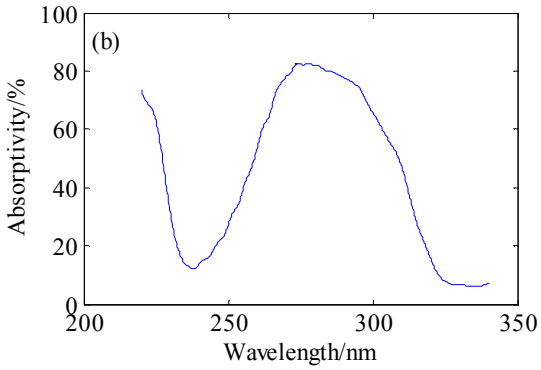
3.1. Determine the Parameter

On the basis of SVM denoising method, the detection result can be divided into two aspects, the first part is the training set, and the second part is the test set. In this article, we totally have 30 samples in data sets, we randomly split the dataset into 25 train sets and 5 test sets. The parameter r of the gaussian radical kernel function has a great relationship with the denoising effect. The penalty factor C is set to 40.18. The parameter r is regulated real-time to get better denoising result. The denoising effects of SO_2 ultraviolet absorption spectrum with different r are shown in Figure 4.

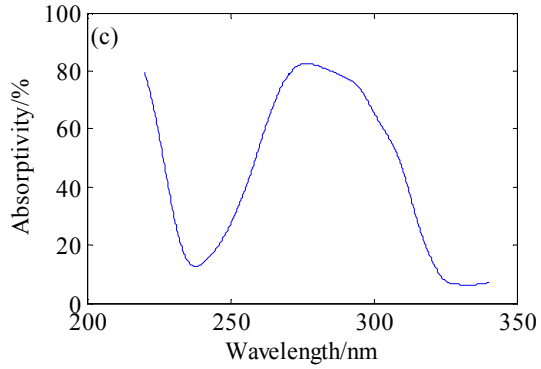
As shown in Figure 4, when r is set to 10, there is a lot of noise; when r is set to 1, the denoising effect has improved; when r is set to 0.1, the denoising is better; when r is set to 0.01, the spectral information has been removed. Therefore, when Gaussian radial basis kernel function is used to denoise, r is set to 0.1.



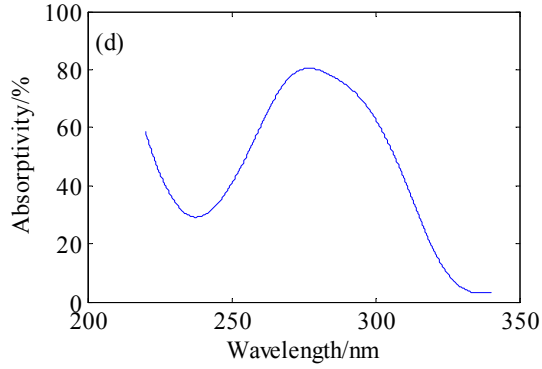
(a) r is set to 10



(b) r is set to 1



(c) r is set to 0.1



(d) r is set to 0.01

Figure 4. *The denoised result when r is change.*

3.2. Denoising Results and Performance Evaluation

The SO_2 ultraviolet absorption spectrum is denoised by wavelet and EMD methods, respectively. Firstly, daubechies11 wavelet transform is used to denoise the SO_2 absorption spectrum. The SO_2 ultraviolet absorption spectrum is decomposed into approximated and three detail coefficients, and shown in Figure 5. The denoising result of wavelet is shown in Figure 6. Secondly, EMD is used to denoise the SO_2 absorption spectrum. The SO_2 ultraviolet absorption spectrum is decomposed into seven IMFs, and shown in Figure 7. The denoising result of EMD is shown in Figure 8. By comparing Figure 4(c), Figure 6 and Figure 8, result shows that SVM has better de-noising effects than wavelet and EMD.

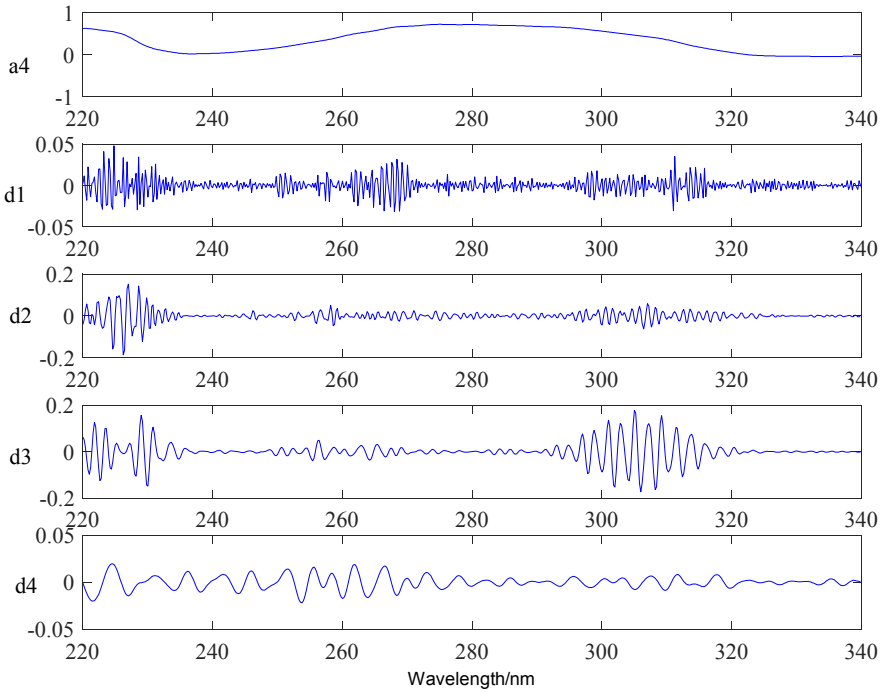


Figure 5. *The wavelet decomposition result of the signal.*

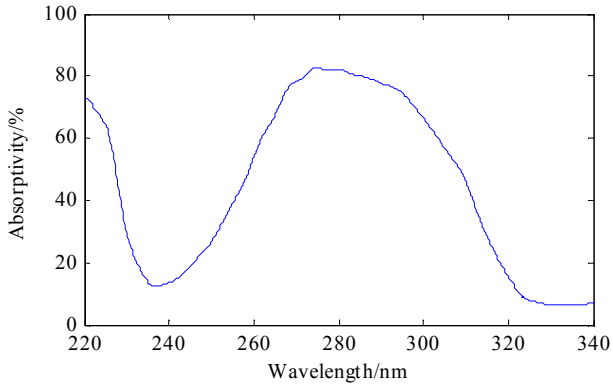


Figure 6. *The wavelet denoised result of the signal.*

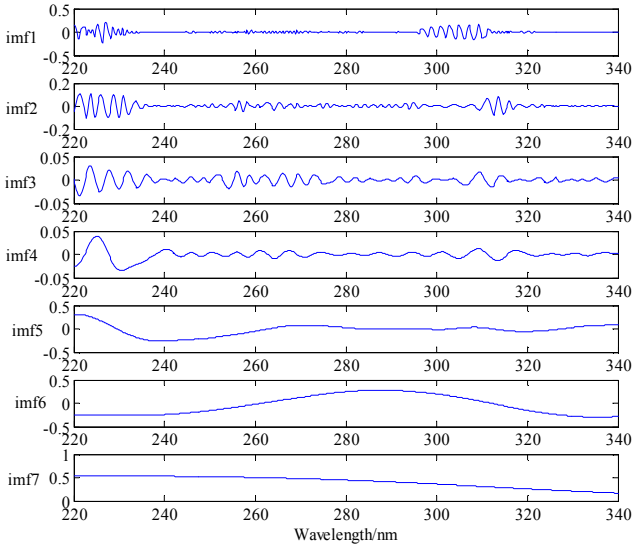


Figure 7. *The EMD decomposition result of the signal.*

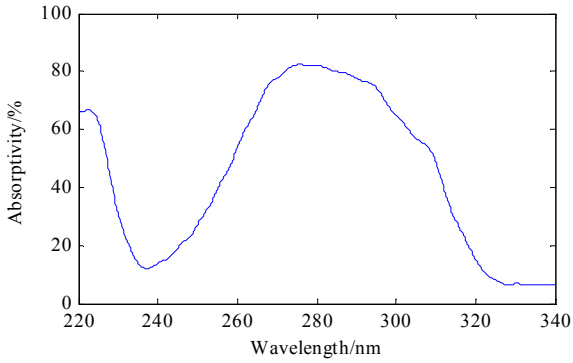


Figure 8. *The EMD denoised result of the signal.*

In this study, to test the accuracy of the wavelet, EMD and SVM algorithms, SNR and MSE are applied for evaluating the denoising accuracy three methods.

$$SNR = 10 \cdot \lg \frac{s}{n} = 10 \cdot \lg \frac{\sum_{i=1}^{N} y_i^2}{\sum_{i=1}^{N} (x_i - y_i)^2} \tag{14}$$

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - x_i)^2 \tag{15}$$

where N is the number of samples, y_i is signal value after denoising at time i, and x_i is original signal value at time i.

SNR indicates the denoising ability of the algorithm. The larger *SNR* value, the better denoising effect. *MSE* indicates that the denoised spectral amplitude is compared with the original spectral amplitude. The smaller *MSE* value, the better denoising effect [26]. The values of the qualitative analysis of three kinds of denoising methods are listed in

Tab.1. From the values of qualitative analysis in Tab.1, it can be seen that the *SNR* of SVM denoising method is larger, and the *MSE* of SVM denoising method is smaller. Therefore, the SVM de-noising method is better than wavelet and EMD.

Table 1. *SNR and MSE values of the three methods.*

method	SVM	wavelet	EMD
SNR/dB	48.9398	33.8470	42.582
MSE	1×10^{-7}	1.05×10^{-6}	1.25×10^{-7}

3.3. System Performance After De-Noising

Denoising results of three denoising methods are applied to the SO_2 detection system. Ten concentrations of SO_2 gas are prepared. The ultraviolet absorption spectrum was obtained by the detection system shown in Figure 2. Ten concentrations of SO_2 spectra were denoised using SVM, wavelet, EMD, respectively. The relationship between the SO_2 concentration and the measured voltage value is fitted after denoising. The fitting curve after SVM denoising is shown in Figure 9, the linear equation is $f(x)=0.7401x+10.72$, the linear correlation coefficient is 0.9971. The fitting curve after wavelet denoising is shown in Figure 10, the linear equation is $f(x)=0.7545x+6.77$, the linear correlation coefficient is 0.9841. The fitting curve after EMD denoising is shown in Figure 11, the linear equation is $f(x)=0.7442x+10.11$, the linear correlation coefficient is 0.9923. Comparing with the wavelet denoising and EMD denoising method, the result showed that SVM denoising method has better de-noising effects for SO_2 detection signal.

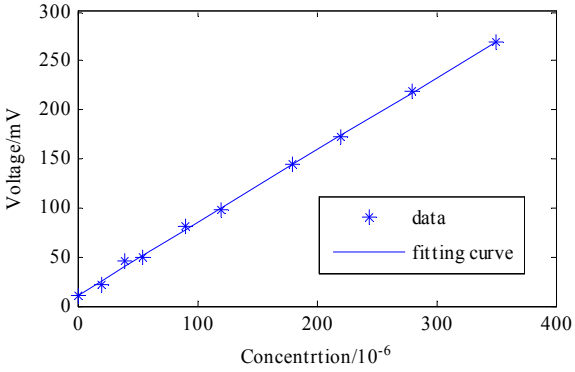


Figure 9. *The fitting curve between the SO_2 concentration and the measured voltage value after SVM denoising.*

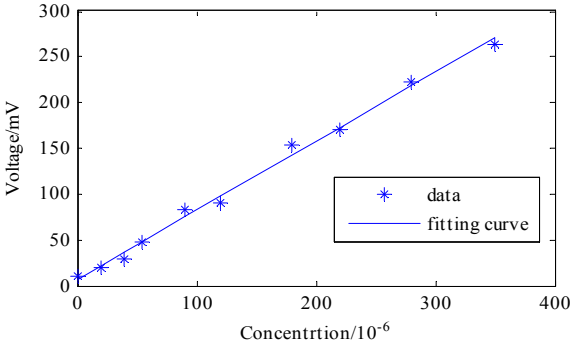


Figure 10. *The fitting curve between the SO_2 concentration and the measured voltage value after wavelet denoising.*

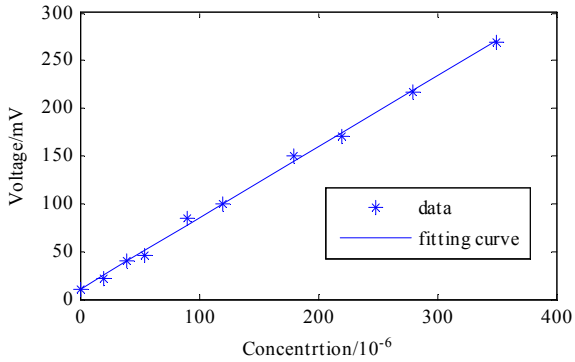


Figure 11. *The fitting curve between the SO_2 concentration and the measured voltage value after EMD denoising.*

4. Conclusions

In order to choose a suitable denoising arithmetic for SO_2 detection system, the SVM method to proposed. The SVM method is applied to the SO_2 ultraviolet absorption spectrum and compared with the wavelet and EMD denoising methods. The experimental results show that *SNR* and *MSE* are improved by SVM denoising. The ultraviolet absorption spectrum denoised by SVM can improve the linear correlation of the SO_2 concentration and the measured voltage value, and make the detection more accurate Those experimental results can prove the effectiveness of the denoising method. The next steps are to apply the SVM denoised method to other gas and to predict gas concentration, this can realize the universality of SVM.

Acknowledgements

The authors gratefully acknowledge the financial support from the National Natural Science Foundation of China (No. 61771419) and Natural Science Foundation of Hebei province of China (No. F2017203220).

References

[1] GUO H B, HUANG S J, Air pollutants and asthma patient visits: indication of source influence. *SCI TOTAL ENVIRON*, 2018, 625: 355-362.

[2] YU Y J, YU Z L, SUN P, et al. Effects of ambient air pollution from municipal solid waste landfill on children's non-specific immunity and respiratory health. *Environmental Pollution*, 2018, 236: 382-390.

[3] ZHAO Y, DUAN L, XING J, et al. Soil acidification in China: Is controlling SO2 emissions enough?. *ENVIRON SCI TECH LET*, 2009, 43(21): 8021-8026.

[4] LIU Y Y, XU X Y, HEN Y, et al. An integrated micro-chip with Ru/Al2O3 /ZnO as sensing material for SO2 detection. *SENSOR ACTUAT B-CHEM*, 2018, 262: 26-34.

[5] LIU W Q, CUI Z C, LIU J G, et al. Measurement of atmospheric trace gases by spectroscopic and chemical techniques. *Chinese Journal of Quantum Electronics*, 2004, 21(2): 202-210.

[6] QI J, DONG X P, ZHENG J D, et al. An algorithm of filtering background noise of optical fiber gas senor. *Chinese Journal of Lasers*, 2011, 38(11): 195-200.

[7] WANG Y T, WANG Y T. Research on de-noising of fluorescence detecting signal based on wavelet transform. *Metrology& Measurement Technique*, 2009, 36(01): 1-2, 5.

[8] BIAN H L, CHEN G F. Anti-aliasing algorithm of nonstationary harmonic signal measurement based on interpolation in frequency domain using short time Fourier transform. *Chinese Journal of Scientific Instrument*, 2008, 29(2): 284-288.

[9] WANG S T, LI M M, LI P, et al. Signal processing method based on empirical mode decomposition in the SO2 concentration monitoring. *Acta Optica Sinica*, 2014, 43(02): 14-19.

[10] ZHAO X M, Patel T H, ZUO M J. Multivariate EMD and full spectrum based condition monitoring for rotating machinery. *Mechanical Systems and Signal Processing*, 2012, 27(1): 712-728.

[11] ZHU B Z, HAN D, WANG P, et al. Forecasting carbon price using empirical mode decomposition and evolutionary least squares support vector regression. *Applied Energy*, 2017, 191:521-530.

[12] HAO Z, ZHAO H L, ZHANG C, *et al*. Estimating winter wheat area based on an SVM and the variable fuzzy set method. *REMOTE SENSING LETTERS*, 2019, 10(4): 343-352.

[13] PEND H B, CHEN G H, CHEN X X, *et al*. Hybrid classification of coal and biomass by laser-induced breakdown spectroscopy combined with K-means and SVM. *PLASMA SCIENCE & TECHNOLOGY*, 2019, 21(3): UNSP 034008.

[14] RAPUR J. S., TIWARI, R.. On-line Time Domain Vibration and Current Signals Based Multi-fault Diagnosis of Centrifugal Pumps Using Support Vector Machines. *JOURNAL OF NONDESTRUCTIVE EVALUATION*, 2019, 38(1): 6.

[15] BAHRAM C. B., Moradi Ehsan, Golshan Mohammad, *et al*. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *SCI TOTAL ENVIRON*, 2019, 651: 2087-2096.

[16] ZHOU Y. L., CHANG F. J., CHANG L. C, *et al*. Multi-output support vector machine for regional multi-step-ahead PM2.5 forecasting. *SCI TOTAL ENVIRON*, 2019, 651: 230-240.

[17] LI C B, LIN S S, XU F Q, *et al*. Short-term wind power prediction based on data mining technology and improved support vector machine method: A case study in Northwest China. *JOURNAL OF CLEANER PRODUCTION*, 2018, 205: 909-922.

[18] Fabelo Himar, Ortega Samuel, Casselden Elizabeth, et al. SVM Optimization for Brain Tumor Identification Using Infrared Spectroscopic Samples. *Sensors (Basel, Switzerland)*, 2018, 18(12).

[19] ROUTRAY S, Ray A K, Mishra C, et al.. Efficient hybrid image denoising scheme based on SVM classification [J]. *Optik-International Journal for Light and Electron Optics*, 2018, 157: 503-511.

[20] SHAHDOOSTI H R, HAZAVEI S M. Combined rippled and total variation image denoising methods using twin support vector machines [J]. *Multimedia Tools & Applications*, 2017, 12: 1-19.

[21] CHEN C Y, LIN M L, ZHANG J. Study on signal filtering based on support vector machine [J]. *Journal of Xi'an Jiaotong University*, 2006, 40(4): 427-431.

[22] JI Q C, LUO X H. Filter design and hardware implementation based on SVM [J]. *Transducer and Microsystem Technologies*, 2018, 37(03): 95-98, 102.

[23] VAPNIK V. The Nature of Statistical Learning Theory [M]. *New York: Springer-Verlag*, 1995.

[24] SMOLA A J, SCHOLKOPF B. A tutorial on support vector regression [J]. *Statistics and Computing*, 2004, 14(3): 199-222.

[25] KEERTHI S S, LIN C J. Asymptotic behaviors of support vector machines with Gaussian kernel [J]. *Neural Computation*, 2003, 15(7): 1667-1689.

[26] LI H, LIN Q Z, WANG Q J, et al. Research on spectrum denoising methods based on the combination of wavelet package transformation and mathematical morphology [J]. *Spectroscopy and Spectral Analysis*, 2010, 30(3): 644-648.