

A Spectral Optimization Method to Enhance Handheld LIBS for T91 Aging Grade Classification Using Lorentzian Profile and Kalman Filter

Zihan Shang,^{a,b} Meirong Dong,^{a,c,*} Weiye Lu,^{a,b} Junbin Cai,^{a,c} Kaijie Bai,^b Feiqiang Tang,^{a,c} Zhichun Li,^{a,c} Shunchun Yao,^{a,c} and Jidong Lu^{a,c}

^a School of Electric Power, South China University of Technology, Guangzhou 510640, P. R. China

^b Guangdong Institute of Special Equipment Inspection and Research Shunde Branch, Foshan 528300, P. R. China

^c Guangdong Province Engineering Research Center of High Efficient and Low Pollution Energy Conversion, Guangzhou, Guangdong 510640, P. R. China

Received: December 04, 2023; *Revised:* January 12, 2024; *Accepted:* January 25, 2024; *Available online:* January 25, 2024.

DOI: 10.46770/AS.2023.297

ABSTRACT: The handheld laser-induced breakdown spectroscopy (LIBS) device has been developed to evaluate the aging grade of T91 heat-resistant steel. However, the reduced size of the device could potentially affect the accuracy of the measurement. To overcome this challenge, a novel spectral preprocessing approach that combines Lorentzian fitting correction and extended Kalman filter is proposed in this work to enable handheld LIBS to reliably evaluate T91 aging grades using support vector machine (SVM) algorithm. The spectral features of T91 samples were measured and revealed that the obtained spectral line broadening is mainly due to Stark broadening. Consequently, the independent spectral peaks in the spectra can be corrected with Lorentzian function to obtain more comprehensive spectral information. Afterwards, the extended Kalman filter is introduced to further diminish the fluctuation in the spectral signal and the device noise during the measurement. The results show that the accuracy of the T91 aging grade classification model based on the new preprocessing method can achieve 94 %. This surpasses the accuracy of the models developed with standard normal variation (SNV) and multi-scatter calibration (MSC) preprocessing methods. The proposed method improves the accuracy of handheld LIBS for T91 aging grade classification and provides a new insight into improving the quality of LIBS spectral data.

INTRODUCTION

With the development of thermal power plants towards large capacity and high parameter, the service life of heat-resistant steel is gaining increasing attention. T91 heat-resistant steel, commonly used as boiler tube steel in supercritical power plants, inevitably begins to age after long-term operation at high temperatures. Aging is primarily characterized by the segregation of alloying elements, changes in microstructure, and deterioration of

mechanical properties.^{1,2} The probability of cracking and tube bursting increases greatly when the aging degree of T91 components is very high, potentially causing economic losses and even casualties in power plants. Therefore, it is crucial to assess the reliability of heat-resistant steel components for ensuring the safe and efficient operation of the plant.

Laser-induced breakdown spectroscopy (LIBS) is a rapid detection technology that utilizes the emission spectra of laser-induced plasma to analyze the elemental composition and content

of samples.³ This technology uses laser pulses to ablate the target sample to generate laser plasma, and qualitative or quantitative analysis of the sample is achieved by analyzing the emission spectral lines of the plasma.^{4,7} It exhibits advantages such as simple sample preparation, non-destructive in-situ measurement, and rapid online analysis, and has been widely applied in various fields.^{8,9} It is worth noting that alterations in the physical and chemical properties of the sample significantly influence LIBS measurements (*i.e.*, the matrix effect), which means matrix effect can be exploited to obtain spectral information about changes in the microstructure of the sample. Rusak *et al.*¹⁰ investigated the change in the Ca/F signal ratio at different depths of the archaeological bones, and found that the signal ratio could be used to distinguish the degree of preservation of the bones. Kang *et al.*¹¹ analyzed the stress distribution in metals using the ratio of ionic to atomic spectral line intensity and confirmed the potential of using LIBS to measure stress in metals.

Our team has proposed the application of the LIBS technique to analyze the failure characteristics of heat-resistant steel.¹² Yao *et al.*¹³ demonstrated a correlation between the intensity ratio of Fe II/Fe I and tensile strength in 12Cr1MoV steels with different pearlitic spheroidization grades. Huang *et al.*¹⁴ analyzed the relationship between the LIBS spectra of T91 and its hardness, and established a T91 hardness calibration model to realize the estimation of T91 mechanical properties. Lu *et al.*^{15,16} investigated the time-spatial evolutions of characteristic spectral lines of T91 at different aging grades and found that during the early stage of plasma expansion, T91 samples with higher aging grades exhibited lower plasma temperatures. Additionally, they also demonstrated the reliability of the T91 aging grade classification model established using the spectral line intensities ratio with the support vector machine (SVM) algorithm. The above studies were conducted based on the laboratory bench-top LIBS system, and they confirmed the feasibility of the LIBS technique for heat-resistant steel failure characterization. To further realize the application in the industrial field, we have also independently developed a portable LIBS device.¹⁷ But the size and weight of the portable device are still limited in practical use.

In recent years, a more portable device, handheld LIBS, has been applied in practical applications such as metallurgical analysis,^{18,19} rock minerals analysis^{20,21} and food safety.^{22,23} However, to reduce the weight and size of the laser and spectrometer, the optical paths in the lightweight laser system are relatively simple and the individual parts of the miniaturized spectrometer have been arranged more compactly. As a result, the laser produces lower energy, which affects its stability. This in turn affects the plasma generation and its stability during the laser breakdown process. And the resolution of the spectrometer is restricted and the detection sensitivity is decreased, ultimately resulting in a lower signal-to-noise ratio and poor-quality spectra.²⁴ To ensure the quality of spectral data, efforts should be made to

improve the accuracy, integrity, and consistency of spectral information. The spectral line shape contains information about the plasma state of the sample, Tang *et al.*²⁵ used a Lorentzian function to fit the K elemental spectral line to reduce the self-absorption effect in LIBS and reflect the real intensity of the line. Yan *et al.*²⁶ improved the characterization ability of spectral information by fusing characteristic information such as spectral intensity, spectral peak area (SPA), and signal-to-noise ratio (SNR). In terms of signal stability, Gurov *et al.*²⁷ found that the high noise-immunity of the extended Kalman filter algorithm can well realize dynamic estimates in parallel fringe background, envelope, and unwrapped fringe phase. Palleschi *et al.*²⁸ discussed the application of the extended Kalman filter in non-linear multivariate LIBS analysis and expanded the effective range of the LIBS calibration curve. It can be concluded that appropriate preprocessing of data can improve the spectral quality and enhance the handheld LIBS analysis capabilities.

In this work, T91 samples with different aging grades are tested by handheld LIBS, and a spectral optimization method based on Lorentzian fitting correction and an extended Kalman filter is proposed to enhance the classification accuracy of T91 samples by the device. To improve the accuracy of spectral data, we maintain the line shape integrity by applying the Lorentzian fitting function to correct the line profile of the spectrum. The issue of unstable spectral data is addressed using the extended Kalman filter. Moreover, the performance is assessed by comparing this method with models constructed through other spectral preprocessing techniques like standard normal variation (SNV) and multi-scatter calibration (MSC).

EXPERIMENTAL

Samples. The aging of T91 heat-resistant steel is accompanied by changes in the metallographic structure, according to China's electric power industry standards, the aging grade of T91 can be divided into five grades based on metallographic characteristics. In order to establish an effective classification model for the handheld LIBS, twelve T91 artificial aging samples (Fig. 1) obtained from high temperature aging tests are analyzed. The aging grade of each sample and the elemental composition of the T91 steel are shown in Table 1 and 2, respectively. The specific information of the aging tests has been described in more detail in our previous work.^{14-17,29}

Fig. 1 T91 artificial aging samples.

Table 1. Aging grade of T91 heat-resistant steel samples

Label	Aging Grade	Label	Aging Grade
T-1	1	T-7	2
T-2	2	T-8	4
T-3	2	T-9	3
T-4	2	T-10	3
T-5	3	T-11	5
T-6	3	T-12	3

Table 2. The elemental composition of T91 heat-resistant steel samples

Element	Values, wt%	Element	Values, wt%
C	0.08-0.12	V	0.18-0.25
Si	0.20-0.50	P	0.002
Mn	0.30-0.60	S	0.01
Cr	8.00-9.50	Ni	0.4
Mo	0.85-1.05		

Fig. 2 Handheld LIBS: (a) Optical path; (b) Actual device.

LIBS setup. In this work, the spectral data of samples are obtained by handheld LIBS (with a mass and dimensions of approximately 2.0 kg and 28×12×30 cm, respectively), as shown in Fig 2. A pulsed Nd: YAG laser (Q-switch, emitting at a wavelength of 1064 nm and a repetition rate ranging from 1 Hz to 10 Hz, 10ns pulse duration and a maximum pulse energy of 15 mJ), a compact spectrometer (Model: AvaSpec-Mini4096CL, CMOS 4096 pixels, spectral range: 290-550 nm, 10 μm slit, spectral resolution: 0.14-0.18 nm) and control system are integrated in the device.

The laser frequency is set to 5 Hz. The integration time of the detector is set to 2 ms and the delay time of the spectrometer is set to 900 ns in the experiment. In each experiment, 3 randomly selected positions are measured with 200 laser pulses per spot on the surface of each sample. Then the data from the three points are averaged as the final spectral data of the sample. A total of 2400 spectral data can be obtained from a single experiment on all twelve samples.

To compare the difference in spectral quality between the handheld LIBS and the bench-top LIBS, a laboratory bench-top LIBS system is also employed for the measurement of T91 samples in this work. A Q-switched Nd: YAG laser (top-hat profile at 1064 nm) and a four-channel optical fiber spectrometer (AvaSpec-2048, 20μm slit, resolution: 0.08-0.10 nm) are assembled in the laboratory bench-top LIBS system, and the laser

and spectrometer are controlled by a digital delay generator (DG535). The laser frequency of the LIBS system is set to 5 Hz and adjusted the laser pulse energy to about 15 mJ. Other parameters such as the spectrometer integration time and the delay time of the spectrometer are set consistent with the parameters in the handheld LIBS.

Methodology

1) *Spectral line broadening.* In a plasma environment, when electronic transitions occur in particles (atoms or ion), the emitted spectral line will have a certain width, which is called line broadening. One of the major contributions to the spectral linewidth in gases at low pressures is the Doppler width.^{30,31} It is well known that Doppler broadening is the broadening of spectral lines due to the Doppler effect, in which the thermal movement of the emitting atoms or ions and resulting Doppler shift of the light that is absorbed or emitted. Based on the observed radiation frequency and the density of emitted particles, the intensity profile of Doppler-broadened spectral line can be represented as a Gaussian profile. The commonly used Gaussian fitting function is shown in Equation (1).

$$Gauss: y = y_0 + \frac{A}{\omega \sqrt{\pi/2}} e^{-2\frac{(x-x_c)^2}{\omega^2}} \quad (1)$$

where y_0 is the intensity of the spectral baseline. A is the peak area. x_c is the central wavelength of the spectral line and ω is the full-width at the half of the maximum (FWHM).

In the denser laser plasma, the main cause of spectral line broadening is Stark broadening, and Stark effect is the principal reason. The Stark effect refers to the phenomenon of changes in the transition energy level differences and emission frequencies of luminescent particles (atoms, ions) in a plasma environment when influenced by an electric field. Therefore, Stark broadening results in the broadening of a spectral line and a shift in the frequency of line center. The Lorentzian profile can be characteristic for this broadening,^{32,33} where the typical function is shown in Equation (2).

$$Lorentz: y = y_0 + \frac{2A}{\pi} \frac{\omega}{4(x-x_c)^2 + \omega^2} \quad (2)$$

2) *Kalman filter.* To guarantee the dependability of spectral analysis results, improving the quality of spectral data has become a crucial step for subsequent analysis and modeling.³⁴ The Kalman filter is extensively employed in the fields of robotics and automation.³⁵ This method has high noise-immunity and can obtain the best estimation of the variable to be determined in the linear case. Therefore, it can be used as a data preprocessing method to improve the signal-to-noise ratio of the data and obtain

data that is closer to the actual system state.³⁶ In engineering applications, without considering the control action, the equation of the linear discrete system can be assumed as:

$$\begin{cases} X_k = \Phi_{k,k-1}X_{k-1} + W_{k-1} \\ Z_k = H_kX_k + V_k \end{cases} \quad (3)$$

where X_k is the estimated state value of the system at time k , $\Phi_{k,k-1}$ is the linear evolution matrix of the system from $k - 1$ to time k , W_{k-1} is the process noise of the system at time $k - 1$; Z_k is the observation value of the system at time k , H_k is the observation matrix of the system at time k , V_k is the observation noise of the system at time k .

The optimal state estimation \hat{X}_k of state X_k can be solved by the following steps:

a) A priori estimate of the current k time is calculated according to the estimated value of the system $k - 1$ time:

$$\hat{X}_{k,k-1} = \Phi_{k,k-1}X_{k-1} \quad (4)$$

$$P_{k,k-1} = \Phi_{k,k-1}P_{k-1}\Phi_{k,k-1}^T + Q_{k-1} \quad (5)$$

where $P_{k,k-1}$ is the covariance matrix of a priori estimate and Q_{k-1} is the covariance matrix of the noise W_{k-1} .

b) The priori estimate is further updated by using the measured value at k time and finally a posterior estimate \hat{X}_k is obtained:

$$K_k = P_{k,k-1}H_k^T(H_kP_{k,k-1}H_k^T + R_k)^{-1} \quad (6)$$

$$P_k = (I - K_kH_k)P_{k,k-1} \quad (7)$$

$$\hat{X}_k = \hat{X}_{k,k-1} + K_k(Z_k - H_k\hat{X}_{k,k-1}) \quad (8)$$

where K_k is called the Kalman gain, R_k is the covariance matrix of the noise V_k .

The Kalman filter is to utilize the observation at each time to obtain the optimal estimate \hat{X}_k of state X_k .

In the case of LIBS spectral signals, the measured signals are subject to fluctuation and uncertainty due to the instability of the excited plasma. Additionally, the acquired signals are often interfered with by noise signals, which tends to degrade the quality of the spectral data and cause the spectral analysis results to deviate further from the real values. The Kalman filter can be used to achieve real-time estimation of the spectral ideal signal and eliminate noise, thereby improving signal quality and stability. To make the Kalman filter more suitable for LIBS, it can be extended according to the characteristics of the LIBS spectra. After applying the extended Kalman filter, the best estimation of the actual spectrum

Fig. 3 Applying the extended Kalman filter method to spectral data.

of the sample can be obtained by taking the actual measured data as the observation value and the ideal spectrum as the estimated state value (Fig. 3).

RESULTS AND DISCUSSION

Spectral line broadening of handheld LIBS device. When the light-emitting particles in the plasma emit photons, they are simultaneously disturbed by the plasma particles surrounding them. This impact leads to the obtained spectral lines to display broadening and line shift.³⁷ This kind of phenomena is also appearing in the spectral information collected by the handheld LIBS device. In various plasma environments, the extent of spectral line disruption varies, with differing key factors influencing spectral line broadening. In general, Stark broadening has been considered as the main broadening source in atmospheric pressure environment. For example, the Stark broadening parameter of the Zn I-1 636.23 nm could be measured by LIBS technique in open air.³⁸ Therefore, to clarify the type of line broadening of handheld LIBS, the spectral lines are firstly fitted and analyzed.

The original spectrum obtained by the handheld LIBS is susceptible to the interference of the continuum background emission. In order to get the actual signal measured by the spectrometer, background subtraction and baseline correction preprocessing are required for the original spectrum. The spectral data collected by the handheld LIBS is shown in Fig. 4. It can be seen that the spectrum can reflect the elemental information of the ablated sample very well. To avoid occasionality in the analysis process, two relatively independent characteristic spectral lines without overlapping peaks of the main constituent elements of T91 are chosen to study the spectral line broadening. In this work, Mo

could be taken as 0.16 nm. Then the fitting functions (1) and (2) could be further simplified to Equation (9) and (10). The fitting results of the two characteristic spectral lines are shown in Fig. 5, Fig. 6, and Table 4.

$$\text{Gauss: } y = \frac{A}{\omega \sqrt{\pi/2}} e^{-2\frac{(x-x_c)^2}{\omega^2}} \quad (9)$$

$$\text{Lorentz: } y = \frac{2A}{\pi} \frac{\omega}{4(x-x_c)^2 + \omega^2} \quad (10)$$

Fig. 4 Handheld LIBS signal of T-10 sample.

Fig. 5 The spectral line fitting of Mo I 357.108 nm: (a) Gaussian fitting; (b) Lorentzian fitting.

Fig. 6 The spectral line fitting of Fe I 495.782 nm: (a) Gaussian fitting; (b) Lorentzian fitting.

Table 3. The Spectral Information of the Analyte Lines

Species	λ (nm)	E_i (cm ⁻¹)	E_k (cm ⁻¹)
Mo I	357.108 nm	25638.567	53636.731
Fe I	495.782 nm	22650.416	42815.855

I 357.108 nm and Fe I 495.782 nm spectral lines (Table 3) from the specimen labeled as T-10 are selected for further analysis.

The data of the two characteristic spectral lines were fitted by Gauss and Lorentz lines respectively after three independent experiments. As the influence of baseline drift was eliminated by baseline correction, the value of y_0 could be approximately taken as 0. Meanwhile, according to the parameters of the handheld LIBS spectrometer, the full-width at the half of the maximum ω

It can be observed that the spectral line fitted with a Lorentzian function is narrower and higher as compared to the spectral line fitted with a Gaussian function. The fitting curve of the Gaussian function approaches zero faster than that of the Lorentzian function near the spectral baseline. As a result, the partial fitting of the Gaussian function on the line wings is inferior to that of the Lorentzian function. For the kernel of the line, a Lorentzian profile can fit the experimental data points well. Alternatively, a Gaussian line profile fitted near the line peak may deviate from the data points, which is particularly evident in the fitting of the first and second sets of experimental data for Fe I 495.782 nm (Fig. 6). As shown in Table 4, when exclusively applying either Gaussian fitting or Lorentzian fitting to the two characteristic spectral lines, the coefficient of determination (R^2) of the fitted spectral lines are very similar. The R^2 is between 0.89 and 0.92 when using Gaussian fitting, while the R^2 takes on values within 0.97 and 0.98 when using Lorentzian fitting. This indicates that if the selected characteristic spectral lines are relatively independent and do not contain any overlapping peak, the same fitting function exhibits an approximate performance on the chosen spectral lines. Besides, the R^2 of Lorentzian fitting is always higher than that of Gaussian fitting for any given characteristic line under the same experimental conditions, which means that the emission lines are greatly affected by the disturbance of the surrounding charged particles during the handheld LIBS measurement. Thus, the spectral line type of the handheld LIBS is closer to the Lorentz profile and the broadening mechanism is mainly Stark broadening.

Correction of handheld LIBS signal. The symmetric Lorentzian profile based on experimental data is regarded as the ideal fitting profile in this study. By comparing the actual spectral data obtained from handheld LIBS and their Lorentzian fitting profiles in Fig. 5 (b) and Fig. 6 (b), it is found that the data points are mainly distributed on both sides of the peak of the ideal fitting profile and there are no data points located at the peak value position of the ideal fitting profile. This results in the intensity of data point at peak position being lower than the ideal intensity, indicating that the actual intensity of the spectral line is underestimated. And Fig. 7 (a) shows that the Lorentzian profile fitted with the data point at peak position as the central wavelength is also quite different from the ideal fitting profile (taking the first experimental data of Fe I 495.782 nm as an instance). As a result, the information contained

Table 4. Profile Fitting Results

Characteristic Spectral Line	Experiment	Gaussian Fitting(R ²)	Lorentzian Fitting(R ²)
Mo I 357.108 nm	1	0.90	0.98
	2	0.90	0.98
	3	0.92	0.97
Fe I 495.782 nm	1	0.92	0.98
	2	0.89	0.97
	3	0.92	0.97

Fig. 7 The Lorentzian profile fitted with the data point as the peak value at Fe I 495.782 nm: (a) before correction; (b) after correction.**Fig. 8** The spectrum after the selected peak was corrected

in the spectral line is partially missing, leading to a decrease in the integrity and accuracy of the spectral data.

There are several reasons why data points are absent at the vertex of the ideal fitted spectral line. The Stark effect could be one possible explanation, whereby the perturbation induced by the electric field of the electrons on the energy levels of a given transition, resulting in the shift of the spectral line shape center. On the other hand, the spectrometer itself has a minimum resolution, which may restrict its ability to precisely measure the spectral value at the peak of the spectral line for each wavelength. Moreover, it also could be the environmental factors affecting the spectral data during the handheld LIBS testing process. To improve the data quality of the handheld LIBS, the experimental data of the characteristic spectral lines are corrected by using the method based on the Lorentzian fitting. The intensity of the data points should approach to the peak value of the ideal fitting profile, and the wavelength of the data points should be adopted as the center wavelength of the spectral lines after correction.

This method selects characteristic spectral lines with high

signal-to-noise ratio and no overlapping peak interference from the spectrum of the sample, and primarily concentrates on rectifying the data points at the peak value for correction. The Equation (10) indicates that for a given spectral line, the peak area A is a certain value. Thereby Equation (10) becomes:

$$A = \frac{y\pi[4(x - x_c)^2 + \omega^2]}{2\omega} \quad (11)$$

Assuming the shape of the ideal fitting profile is always symmetric and the spectral data points can be uniformly distributed along the ideal fitting line. Ideally, the corrected data point at the peak value is recorded as (x_c, y_c) and the spectral intensity y_{c-1} and y_{c+1} of the adjacent data points x_{c-1} and x_{c+1} on either side of the central wavelength of the spectral line x_c are equal. And an ideal Lorentzian profile can pass through both y_{c-1} and y_{c+1} simultaneously at this point.

However, for the actual spectral data value y_{c-1} is not equal to y_{c+1} in most cases, it is difficult for y_{c-1} and y_{c+1} to satisfy the ideal fitting profile distribution when x_c is constant. For this reason, by taking (x_{c-1}, y_{c-1}) and (x_{c+1}, y_{c+1}) into Equation (11) respectively with x_c as the central wavelength of the spectral line, there will be two Lorentzian fitting profile with peak areas noted as A_1 and A_2 . Then the average area A_0 of A_1 and A_2 can be taken as the area of the ideal Lorentzian fitting, which can make the closest match to y_{c-1} and y_{c+1} . The correction intensity y'_c of the data point at x_c can be determined by substituting x_c and A_0 into Equation (10).

$$A_0 = \frac{A_1 + A_2}{2} \quad (12)$$

$$y'_c = \frac{2A_0}{\pi\omega} \quad (13)$$

The correction of Fe I 495.782 nm spectral line is shown in Fig. 7 (b). The intensity of the corrected data point is equivalent to the peak value of the ideal Lorentzian fitting. Compared with Fig. 7 (a), the fitted line obtained based on the corrected data point exhibits a closer resemblance to the ideal fitting profile, which ensures the integrity and accuracy of the information transmitted by the spectral line. Furthermore, spectral lines that are relatively independent and have no overlapping peaks, like Fe I 495.782 nm, can be selected for correction. The correction of the selected spectral lines in the averaged spectra (Fig. 8) resulted in the improvement of the line intensity, and the relative relationships between the spectral lines became more distinct.

Fig. 9 The fluctuation of spectral line signal measured by different devices: (a) Handheld LIBS device; (b) LIBS system.

Extended Kalman filter. Spectral correction can improve the integrity and accuracy of spectral data, but has not yet to address issue of spectra stability. In order to analyze the spectral fluctuations of the handheld LIBS, we employed the T-10 sample as the experimental object and compared its spectral data with that obtained from laboratory bench-top LIBS system. The spectral data obtained are normalized by dividing it with the spectral intensity at Fe I 438.397 nm, which maps all data to the same scale. This can effectively eliminate the influence of laser energy fluctuation. The spectral intensities of Mo I 357.108 nm and Fe I 495.782 nm were chosen to compare the fluctuations of the two devices, as shown in Fig. 9.

From the line profile, the spectral line shapes obtained from the laboratory bench-top LIBS system are relatively complete, whereas the characteristics of the handheld LIBS line shapes at the peak value are not well reflected. The stability of the spectral signal of the device can be represented by calculating the standard deviation (STD) of the selected spectral line intensity under the experimental measurement. And the STD of the spectral data at Mo I 357.108 nm and Fe I 495.782 nm measured by the bench-top LIBS system are 0.029 and 0.032, respectively, which are much lower than the 0.062 and 0.068 of the handheld LIBS. Therefore, it can be inferred that the spectral signals of the characteristic spectral lines obtained by the laboratory bench-top

LIBS system are more stable than the handheld device. Considering the influence of laser energy fluctuation on the spectral intensity during measurement has been mitigated by normalization, the fluctuation in signals can only be attributed to performance differences between the devices. This implies the need to use a data processing method to reduce the fluctuations of spectral data for the same aging grade, improve the quality of spectral data and enhance the clustering of data across different aging grades.

The Kalman filter method has been extended to stabilize the spectral signal of the handheld LIBS. Based on the Lorentzian fitting correction method in Section 3.2, the estimated state value of the system is determined using the corrected mean spectra of the same aging grade while the observation value of the system consists of spectral data collected by the handheld LIBS spectrometer. As the target spectrum is static for estimation, the $\Phi_{k,k-1}$ matrix becomes an identity matrix in Equation (4) and (5). In practice, H_k is replaced by the Jacobian of $f(X)$.

$$\hat{X}_{k,k-1} = f(X) \quad (14)$$

$$P_{k,k-1} = P_{k-1} + Q_{k-1} \quad (15)$$

$$H_k = \frac{\partial f(X)|_{X_{k-1}}}{\partial X} \quad (16)$$

where $f(X) = X_{k-1}$.

Fig. 10 Spectral optimization at Mo I 357.108 nm and Fe I 495.782 nm: (a) Spectral peak value correction processing; (b) Extended Kalman filter processing.

During the measurement process, it is crucial to note that the spectral data obtained from each laser excitation is independent and unequal due to the laser plasma instability. The inconsistencies in plasma formation following each laser shot, along with instrumental effects throughout the measurement process, are considered as system noise. Then Equations (6) to (8) can be simplified as:

$$K_k = P_{k,k-1}(P_{k,k-1} + R_k)^{-1} \quad (17)$$

$$P_k = (I - K_k)P_{k,k-1} \quad (18)$$

$$\hat{X}_k = \hat{X}_{k,k-1} + K_k(Z_k - \hat{X}_{k,k-1}) \quad (19)$$

Figure 10 shows the optimization process of the handheld LIBS spectral data at Mo I 357.108 nm and Fe I 495.782 nm. After the spectral peak value correction (Fig. 10 (a)), the peak intensity that may have been underestimated was compensated, resulting in a more complete spectral line profile. But the fluctuation of the corrected spectrum was not improved, the STD of the two characteristic spectral line data remained at a high level. Subsequently, after the extended Kalman filter optimization (Fig. 10 (b)), the fluctuation is significantly reduced and the STD of the characteristic spectral line data is also decreased to 0.017 and

0.025, which is equivalent to the STD level of the spectral data obtained by the laboratory bench-top LIBS system. This suggests that the application of extended Kalman filter ensures the consistency of spectral data obtained by handheld LIBS. Thus, a new spectral preprocessing method has been developed by combining the Lorentzian fitting correction with extended Kalman filtering to optimize handheld LIBS spectral data.

To ensure that the modeling is supported by sufficient and reliable data, spectral data from two samples at each aging grade should be selected for model training. Thus, T-2, T-4, T-10 and T-12 were selected as samples for aging grades 2 and 3, while only one sample could be selected for aging grades 1, 4 and 5 respectively (Table 1). In order to balance the amount of data for each aging grade category in the training and to prevent the model from degrading accuracy and generalization performance caused by an imbalanced dataset³⁹, it is necessary to resample the spectral data for samples at aging grades 1, 4, and 5. Under these circumstances, two additional experiments are carried out on T-1, T-8, and T-11 to balance the sample distribution in each category. The SVM algorithm maps the data to a high dimensional space using a kernel function and classifies the data by finding the optimal hyperplane in that space. Furthermore, the reliability of SVM in T91 classification has been proved.¹⁶ The dataset is divided into 80% as a training set and 20% as a testing set, and

Table 5. The performance of the models after different preprocessing methods

Preprocessing method	Training accuracy (%)	Test accuracy (%)	Macro precision (%)	Macro recall (%)
Original spectrum	95.60	68.80	68.38	68.76
SNV	100.00	70.13	70.27	70.15
MSC	99.20	88.53	89.15	88.48
Lorentzian fitting and Kalman	99.07	94.13	94.29	94.19

after separate preprocessing, the training set data is modeled using SVM algorithm. And to enable the model to extract sufficient information from spectral data and maintain great generalization ability, the establishment of T91 aging grade classification model is based on full-spectrum data.

In addition to the spectral optimization method proposed other commonly used preprocessing methods such as multi-scatter calibration (MSC)⁴⁰ and standard normal variation (SNV)⁴¹ were compared in this work. The MSC corrects the spectra by fitting each spectral line and the average spectrum with a linear regression equation. And the SNV corrects the spectra by assuming that the spectral values at each wavelength should satisfy a normal distribution. The performance of the models established by three different preprocessing methods combined with SVM algorithm is presented in Table 5.

Considering the classification accuracy of the training set and the test set, the accuracy of all data preprocessing models has improved. Among them, the accuracy of the model pretreated by the extended Kalman filter method on the test set has reached 94.13%, which is the highest. Although the SNV method achieved a perfect accuracy of 100% on the training set, its accuracy on the test set was only 70.13%, indicating that the training model was overfitting. The possible reason is that SNV is relatively sensitive to outliers, which can lead to overfitting of the data and limit the ability to minimize spectral fluctuations. Besides, the MSC method assumes a linear relationship between each spectrum and their average spectrum (ideal spectrum), without considering non-linear cases. In contrast, the extended Kalman filter method incorporates the correlation between each spectral line intensity and its ideal spectral line intensity in the spectrum, enabling the optimal dynamic estimation of the ideal spectrum for each individual spectrum. From the macro precision and macro recall of each model, the extended Kalman filtering method has also achieved over 94%, demonstrating that the robustness of the model pretreated by the extended Kalman filter method is guaranteed.

CONCLUSION

To improve the spectral quality of the handheld LIBS device and the classification accuracy of the T91 aging grade, we propose a novel approach for optimizing the spectra using Lorentzian fitting

correction and the extended Kalman filter. It has been verified that the predominant mechanism leading to spectral line broadening in handheld LIBS is Stark broadening, and its line profile can be well-fitted by the Lorentzian function. The data points with underestimated intensities at the spectral peak value are then corrected using the Lorentzian fitting, which ensures the integrity and accuracy of the spectral information. To address the issue of spectral instability from the performance of the handheld LIBS, we employ the extended Kalman filter method to improve the consistency of spectral data and the quality of the spectrum. The findings illustrate that the aging grade classification model established by the proposed spectral optimization method for T91 heat-resistant steel exhibits outstanding accuracy and generalization ability. The accuracy of the model on the test set can reach up to 94.13%, which guarantees the application of handheld LIBS for T91 aging grade detection. The method can be used as a data preprocessing technique to facilitate high-quality spectral data acquisition for handheld LIBS and to enhance its ability to manage the challenges that may be encountered during industrial field testing.

AUTHOR INFORMATION



Meirong Dong received her Ph.D. degree at South China University of Technology (SCUT) in the research group “Energy Clean Conversion and System Optimization” headed by Prof. Jidong Lu. She worked as a visiting student under the supervision of senior scientist Dr. Richard E. Russo at Lawrence Berkeley National Laboratory (LBNL) from Sept. 2011 to Aug. 2013. Her research is mainly focused on Advanced optical measurement technologies (LIBS, TDLAS, and PLIF), Intelligent and optimized operation of energy conversion system and equipment, Conversion mechanism of energy substances. She has been working as an editorial board member of *Atomic Spectroscopy*. Mei-Rong Dong is author or co-author of over 80 articles published in peer-reviewed scientific journals, with an h-index of 20 (Web of Science).

Corresponding Author

*M. R. Dong

Email address: epdongmr@scut.edu.cn

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The research was supported by the National Natural Science Foundation of China (U22B20119), and Guangdong Basic and Applied Basic Research Foundation (2022A1515010709). We also acknowledge the support from the Fundamental Research Funds for the Central Universities (2022ZFJH04) and Guangdong Province Key Laboratory of Efficient and Clean Energy Utilization (2013A061401005).

REFERENCES

1. V. Lok, T. G. Le, J. M. Yu, Y. W. Ma, V. P. Nguyen, S. T. Choi, and K. B. Yoon, *J. Mech. Sci. Technol.*, 2020, **34**, 3283–3293. <https://doi.org/10.1007/s12206-020-0720-4>
2. C. G. Panait, W. Bendick, A. Fuchsmann, A. F. Gourgues-Lorenzon, and J. Besson, *Int. J. Pressure Vessels Pip.*, 2010, **87**, 326–335. <https://doi.org/10.1016/j.ijpvp.2010.03.017>
3. D. A. Rusak, B. C. Castle, B. W. Smith, and J. D. Winefordner, *Trac-Trends Anal. Chem.*, 1998, **17**, 453–461. [https://doi.org/10.1016/s0165-9936\(98\)00058-2](https://doi.org/10.1016/s0165-9936(98)00058-2)
4. G. Rao, M. Dong, W. Nie, X. Lin, Y. Liang, and J. Lu, *Atom. Spectrosc.*, 2021, **42**, 349–358. <https://doi.org/10.46770/as.2021.099>
5. Y. Zhang, M. Dong, L. Cheng, L. Wei, J. Cai, and J. Lu, *J. Anal. At. Spectrom.*, 2020, **35**, 810–818. <https://doi.org/10.1039/C9JA00429G>
6. W. Ma, Z. Yu, Z. Lu, Q. Ma, and S. Yao, *Atom. Spectrosc.*, 2023, **44**, 160–168. <https://doi.org/10.46770/as.2023.102>
7. K. K. Ayyalomasayajula, Y.-Y. Fang, J. P. Singh, D. L. McIntyre, and J. Jain, *Appl. Optics*, 2012, **51**, B149–B154. <https://doi.org/10.1364/ao.51.00b149>
8. R. Noll, C. Fricke-Begemann, M. Brunk, S. Connemann, C. Meinhardt, M. Scharun, V. Sturm, J. Makowe, and C. Gehlen, *Spectrochim. Acta B*, 2014, **93**, 41–51. <https://doi.org/10.1016/j.sab.2014.02.001>
9. Y. Chu, Z. Zhang, Q. He, F. Chen, Z. Sheng, D. Zhang, H. Jin, F. Jiang, and L. Guo, *J. Adv. Res.*, 2020, **24**, 353–361. <https://doi.org/10.1016/j.jare.2020.05.001>
10. D. A. Rusak, R. M. Marsico, and B. L. Taroli, *Appl. Spectrosc.*, 2011, **65**, 1193–1196. <https://doi.org/10.1366/11-06364>
11. D. Kang, S. Yoon, and J. Kim, *J. Micromech. Microeng.*, 2019, **29**, 125008. <https://doi.org/10.1088/1361-6439/ab4b7f>
12. M. Dong, J. Cai, S. Yao, and J. Lu, Application of LIBS for the Failure Characteristics Prediction of Heat-Resistant Steel in Laser Induced Breakdown Spectroscopy (LIBS). 2023. <https://doi.org/10.1002/9781119758396.ch45>
13. S. Yao, M. Dong, J. Lu, J. Li, and X. Dong, *Laser Phys.*, 2013, **23**, 125702. <https://doi.org/10.1088/1054-660X/23/12/125702>
14. J. Huang, M. Dong, S. Lu, W. Li, J. Lu, C. Liu, and J. H. Yoo, *J. Anal. At. Spectrom.*, 2018, **33**, 720–729. <https://doi.org/10.1039/c8ja00030a>
15. S. Lu, M. Dong, Q. Chen, J. Huang, Y. Yu, and J. Lu, *Spectrochim. Acta B*, 2019, **151**, 1–11. <https://doi.org/10.1016/j.sab.2018.11.004>
16. S. Lu, M. Dong, J. Huang, W. Li, J. Lu, and J. Li, *Spectrochim. Acta B*, 2018, **140**, 35–43. <https://doi.org/10.1016/j.sab.2017.11.005>
17. J. Cai, M. Dong, Y. Zhang, Y. Chen, H. Chen, Y. Liang, W. Li, and J. Lu, *Atom. Spectrosc.*, 2021, **42**, 43–50. <https://doi.org/10.46770/as.2020.217>
18. C. Fabre, N. E. Ourti, C. Ballouard, J. Mercadier, and J. Cauzid, *J. Geochem. Explor.*, 2022, **236**, 106979. <https://doi.org/10.1016/j.gexplo.2022.106979>
19. L. Brooks and G. Gaustad, *J. Sust. Metall.*, 2021, **7**, 732–754. <https://doi.org/10.1007/s40831-021-00361-3>
20. R. S. Harmon, D. Khashchevskaya, M. Morency, L. A. Owen, M. Jennings, J. R. Knott, and J. M. Dortch, *Molecules*, 2021, **26**, 5200. <https://doi.org/10.3390/molecules26175200>
21. R. S. Harmon, R. R. Hark, C. S. Throckmorton, E. C. Rankey, M. A. Wise, A. M. Somers, and L. M. Collins, *Geostand. Geoanal. Res.*, 2017, **41**, 563–584. <https://doi.org/10.1111/ggr.12175>
22. W. H. Yan, X. Y. Yang, X. Geng, L. S. Wang, L. Lu, Y. Tian, Y. Li, and H. Lin, *Spectrosc. Spectr. Anal.*, 2022, **42**, 3714–3718. DOI: 10.3964/j.issn.1000-0593(2022)12-3714-05
23. X. Wu, S. Shin, C. Gondhalekar, V. Patsekina, E. Bae, J. P. Robinson, and B. Rajwa, *Foods*, 2023, **12**, 402. <https://doi.org/10.3390/foods12020402>
24. G. S. Senesi, R. S. Harmon, and R. R. Hark, *Spectrochim. Acta B*, 2021, **175**, 106013. <https://doi.org/10.1016/j.sab.2020.106013>
25. Y. Tang, L. Guo, S. Tang, Y. Chu, Q. Zeng, X. Zeng, J. Duan, and Y. Lu, *Appl. Optics*, 2018, **57**, 6451–6455. <https://doi.org/10.1364/ao.57.006451>
26. J. Yan, F. Fan, Q. Li, Z. Tang, W. Chen, W. Jiang, Q. Li, L. Yu, and X. Li, *Spectrochim. Acta B*, 2022, **190**, 106375. <https://doi.org/10.1016/j.sab.2022.106375>
27. I. Gurov, P. Hlubina, M. Taratin, and A. Zakharov, in Conference on Optical Measurement Systems for Industrial Inspection IV. Munich, GERMANY, 2005, vol. 5856, pp 274–280. <https://doi.org/10.1117/12.612713>
28. A. Palleschi and V. Palleschi, *Spectrochim. Acta B*, 2018, **149**, 271–275. <https://doi.org/10.1016/j.sab.2018.09.003>
29. J. Li, J. Lu, Y. Dai, M. Dong, W. Zhong, and S. Yao, *Appl. Surf. Sci.*, 2015, **346**, 302–310. <https://doi.org/10.1016/j.apsusc.2015.03.186>
30. I. B. Gornushkin, L. A. King, B. W. Smith, N. Omenetto, and J. D. Winefordner, *Spectrochim. Acta B*, 1999, **54**, 1207–1217. [https://doi.org/10.1016/S0584-8547\(99\)00064-6](https://doi.org/10.1016/S0584-8547(99)00064-6)
31. I. I. Sobel'man and L. A. Vainshtein, Excitation of atoms and broadening of spectral lines., 1995. <https://link.springer.com/book/9783642965586>
32. X. Z. Zhao, L. J. Shen, T. X. Lu, and K. Niemax, *Appl. Phys. B*, 1992, **55**, 327–330. <https://doi.org/10.1007/bf00333075>
33. Hans R. Griem, Principles of Plasma Spectroscopy. 2005. <https://ui.adsabs.harvard.edu/abs/2005ppps..book.....G/abstract>
34. L. Li, X. Liu, F. Yang, W. Xu, J. Wang, and R. Shu, *Spectrochim. Acta B*, 2021, **180**, 106183. <https://doi.org/10.1016/j.sab.2021.106183>

35. P. Zarchan, *Progress in astronautics and aeronautics: fundamentals of Kalman filtering: a practical approach*, 2005. <https://iaac.technion.ac.il/workshops/2010/KFhandouts/LectKF1.pdf>
36. S. S. G. Ram, *International Conference on Electrical, Computer and Communication Technologies (ICECCT)*. 2017. <https://doi.org/10.1109/ICECCT.2017.8117921>
37. S. Messaoud Aberkane, A. Safi, A. Botto, B. Campanella, S. Legnaioli, F. Poggialini, S. Raneri, F. Rezaei, and V. Palleschi, *Appl. Sci.*, 2020, **10**, 4973. <https://doi.org/10.3390/app10144973>
38. A. M. E. Sherbini, A.-N. Aboulfotouh, F. Rashid, S. H. Allam, A. M. Al-Kaoud, A. E. Dakrouri, and T. M. E. Sherbini, *Natural Science*, 2013, **5**, 501–507. <https://doi.org/10.4236/ns.2013.54063>
39. S. Matwin and J. Miłniczuk, *Challenges in Computational Statistics and Data Mining*, Springer International Publishing: Cham, 2016. <https://link.springer.com/content/pdf/10.1007/978-3-319-18781-5.pdf>
40. T. T. Gong, Y. Tian, Q. Chen, B. Y. Xue, F. Z. Huang, L. T. Wang, and Y. Li, *Spectrosc. Spectr. Anal.*, 2020, **40**, 1207–1213. DOI: 10.3964/j.issn.1000-0593(2020)04-1207-07
41. D. Syvilay, N. Wilkie-Chancellor, B. Trichereau, A. Texier, L. Martinez, S. Serfaty, and V. Detalle, *Spectrochim. Acta B*, 2015, **114**, 38–45. <https://doi.org/10.1016/j.sab.2015.09.022>
-