Classification of Chinese Tea Leaves by Laser-induced Breakdown Spectroscopy Combined with Discriminant Analysis Method

Jinmei Wang,[a,c] Peichao Zheng, [*b,c] Hongdi Liu,[c] and Liang Fang[*a]

[a] College of Physics, Chongqing University, Chongqing 400044, P. R. China.

[b] State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, Chongqing 400044, P. R. China.

[c] College of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing Municipal Level Key Laboratory of Photoelectronic Information Sensing and Transmitting Technology, Chongqing 400065, P. R. China.

Abstract: Six kinds of tea leaves, including Longjing Green Tea, Mengding Huangya, White tea, Tie Guanyin, Wuyi black tea and Pu'er tea, were analyzed and identified by laser-induced breakdown spectroscopy (LIBS) combined with discriminant analysis (DA) method. The spectral lines intensities of Mg, Mn, Ca, Al, Fe, K, CN (0-0) and C_2 (0-0) reference to C I 247.86 nm were selected as the analysis indexes according to the differences of LIBS spectra. The results show that 294 samples are identified correctly from 300 training set, and the average correct identification rate is 98%. In addition, 286 samples from 300 testing set are correctly identified, and the average correct identification rate is 95.33%. The misclassification mainly occurred among Mengding huangya tea, White tea and Black tea due to little difference of the line intensities of Mg, Mn, Ca, Al and C_2 (0-0) among these tea leaves. These research results provide available reference method for the identification of the tea leaves.

Keywords: Chinese Tea Leaves; Laser-induced Breakdown Spectroscopy; Discriminant Analysis Method

1. Introduction

The tea leaves originating from China, has a variety of beneficial components for human body, [1]. In addition, with the improvement of people's living standard, the overall living quality of the residents is higher and higher. How to enjoy good quality tea reasonably has attracted more and more attention. However, in recent years, the issue of food quality and safety become more and more serious, which pushes people to take care of the quality and safety of food, and tea leaves are no exception. In China, there are various kinds of tea leaves, and the classification standards vary. Many factors such as the processing technique, the original place and the storage method affect the quality of the tea leaves, which leads to the uneven prices between them. [2] For this reason, some criminals use the original tea instead of unoriginal tea leaves or take the inferior tea leaf instead of the first-class tea leaf to reap sudden huge profits. Thus it is necessary to find out rapid, accurate and reliable tea identification methods, and guarantee the quality of tea.

Nowadays, several analytical methods have been introduced to analyze the tea leaves, for example, atomic absorption spectroscopy (AAS)[3], inductively coupled plasma-mass spectroscopy (ICP-MS)[4], inductively coupled plasma-atomic emission spectroscopy (ICP-AES)[5], gas chromatography (GC) [6], high performance liquid chromatography (HPLC) [7], atomic fluorescence

[*] Correspongding author
[*] Correspongding author

spectroscopy (AFS) [8] and near-infrared spectroscopy (NIRS) [9]. However, these methods mentioned above have some inherent defects more or less. Some of these techniques cost a lot of time in sample pretreatment. What is worse, some new impurities may be introduced in the process, and may seriously affect the detection results. Besides, some of these detection methods' process are relatively complex, time-consuming, and beyond the ability to achieve remote and on-line detection. All these drawbacks lead to the disability to meet the detection requirements of some special fields. It is necessary to investigate rapid, accurate and reliable tea detection methods.

In recent years, laser-induced breakdown spectroscopy (LIBS), a versatile elemental analytical method, has been employed for scientific investigation and technical application in a broad area of topical fields [10-11], which show great advantages in material composition detection over conventional elemental analytical methods. Nowadays, some researchers tried to combine LIBS technique with ripe statistical methods and algorithm for the detection and identification of substances, and have made some achievements. Moncayo et al. developed a fast and minimally destructive method based on LIBS and Neural Networks (NN) to classify and discriminate the human bones and teeth fragments. The methodology can be useful in Disaster Victim Identification (DVI) tasks. The elemental compositions of bone and teeth samples provided enough information to achieve a correct discrimination and reassembling of different human remains [12]. Liang et al. employed LIBS technique combined SVM to classify the classes of steels, and most of the ambiguous data and control the computation cost within an acceptable range are effectively distinguished[13]. Unnikrishnan et al. applied LIBS technique to detect four widely used plastics, and use Principal Component Analysis (PCA) of the LIBS spectra to classify the plastics, and achieve higher accurate results of classification [14]. Yu et al. also developed a method combined LIBS technique and support vector algorithm for the identification of eleven kinds of plastic, and achieved the average correction identification rate of 98.73% [15].

Discriminant analysis is a method used in statistics, pattern recognition and machine learning to find a combination of features according to a certain amount of samples of a grouping variable and the corresponding other multivariate variables of known information to characterize or separates two or more classes of objects or events. Since the discriminant analysis method is a supervised learning algorithm, it is easy to realize and the results are repeatable. In this work, the LIBS technique was employed to detect components in the Chinese tea leaves. Based on the LIBS spectra for tea leaves, the discriminant analysis method was introduced to discriminate 6 kinds of tea leaves.

2. Experimental

2.1 Instrumental setup

The experimental setup in this work is similar to the one used in the previous study [16], which is illustrated in Fig.1. A pulsed Nd: YAG laser (Quantel, Ultra 100) was used as excitation light source, and the laser operated at fundamental wavelength of 1064 nm with pulse width of 5.82 ns. The maximum output laser energy is 100 mJ and the far-field divergence angle of the laser beam is less than 1 mrad. The energy of the laser could be adjusted by changing the Q delay time, which could be evaluated using a laser power meter. The laser was focused on the surface of the target with a 100 mm focal length plano-convex quartz lens to produce intense plasma. The sample was mounted on a three dimensional translation stage, and kept moving to provide a fresh surface after each laser ablation. A homemade collection device consist of a plano-convex quartz lens (f=75 mm) and another biconvex quartz lens (f=30 mm) inclined at angle 45° with the direction of the incident laser beam was used to collected the light signal emitted from the produced plasma, and then coupled the light into a 2 m long multimode silica fiber, and the numerical aperture of fiber bundle is 0.22. Finally, the light was transmitted through the fiber to the entrance of a computerized Czerny-Turner spectrograph (Andor Model SR-750-A) equipped with three ruled gratings: 2400, 1200 and 300 grooves/mm, which were interchangeable under computer control, providing 0.02nm, 0.04nm and 0.18 nm resolution spectra in the wavelength range from 200 nm to 900 nm. The ruled gratings with 300 grooves/mm is blazed at 300nm, offering a wavelength response from 200nm to 450nm and the others is blazed at 500nm which offer the response range from 250 to 900 nm. A gateable and intensified charge-coupled device (ICCD) camera (Andor, istar, 2048×512 pixels) coupled to the spectrometer was used for detecting the dispersed plasma emission light. Then the acquired spectral data were transferred to the personal computer with the high speed USB connection cable. All function of the ICCD camera and the spectrograph were controlled by Andor SOLIS software provided by the manufacturer. In order to reduce the noise, the ICCD camera was cooled to -15 ℃ by Peltier cooler. Thirty pulses were accumulated to obtain one spectrum in order to reduce the system error under the same condition. All experimental operations were performed in open air at atmospheric pressure.

The parameters of laser and spectrometer are important for LIBS technique. In order to reduce the interference from the noise signal and obtain the significant and stable signal, a similar process of parameters optimization for the laser and spectrometer were carried out based on previous studies in our laboratory[17], and the conditions of 1.2 μs ICCD delay time, 4 μs ICCD gate width and 50 mJ laser pulse energy were selected for the experiments.

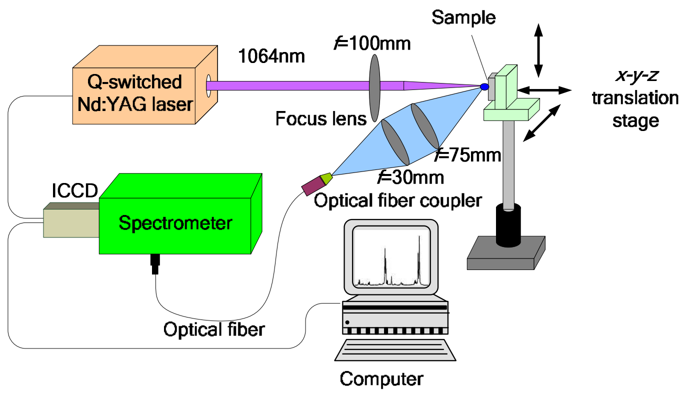

Fig.1 Experimental setup of LIBS for identification of different Chinese tea leaves

2.2 Sample preparation

Six kinds of representative Chinese tea leaves were selected to carry out the experiments, which are Longjing Green Tea(LJ), Mengding Huangya(MH), White tea(WH), Tie Guanyin(TG), Wuyi black tea(WB) and Pu'er tea(PE), respectively. According to the previous experimental experiences, the physical proprieties of the samples such as the dryness, lapping uniformity and density impact on the spectral signal very much. [17] To greatly reduce the influences on the results, some necessary pretreatment of the tea leaves sample need to be done before the determination. At the outset, these six kinds of tea leaves were dried by the electric air blowing drying box for five hours under 80℃, and then grinded with agate mortar for one hour to obtain uniform and fine powder. Finally, these six kinds of powders were compressed to tablets (ϕ16× 3 mm) using a mechanical compression machine at 40 MPa for 2 min.

2.3 Experimental methods-Discriminant analysis method

The objects of this work is to identify the tea leaves by LIBS under the assistance of the Discriminant Analysis Method, and the SPSS statistics software package was used to analyze the experiment. The software manufacture chooses Fisher discriminant and Bayes discriminant analysis method to analyze the data, and the output is mainly presented by Fisher discriminant analysis method.

Fisher Discriminant Analysis (FDA) is one of the most popular dimensionality reduction methods. As a kind of supervised algorithm, it seeks an embedding transformation on which the data examples of different classes are far from each other while requiring data examples of the same class to be close to each other. [18]

Bayes statistics is the basis of Bayes discriminant analysis, and the primary idea of Bayes statistics is to assume there is certain cognition of the entire population before sampling, the prior probability distribution can be used to describe this kind of cognition. Then modify the priori knowledge based on the samples, and get the posterior probability distribution. Finally, make a variety of statistical inference based on the posterior probability distribution.

3. Results and discussion

In the present experiment, the pulsed laser was used to ablate the tablet samples to produce plasma, and the emission spectra were measured in the ultraviolet and visible regions from 200 to 800 nm. Fig. 2 shows the LIBS spectrum of Longjing Green tea leaf, and the elemental compositions of the tea leaves were identified and marked according to NIST Atomic Spectra Database. [19] From the LIBS spectrum, the characteristic spectral lines of Ca, Fe, K, Al and Na etc. metal elements were observed, besides some nonmetallic elements such as C, H, O, N and Si can be observed too. In addition, some molecular spectra such as CN and C_2 appear in LIBS spectrum. The CN emissions are coming from excited carbon atoms from the tea leaves matrix and excited nitrogen atoms from the ambient air following the recombining process ($C_2+N_2\rightarrow 2CN$).[20] In order to catch the fingerprint information of these tea leaves, and distinguish different types of tea leaves correctly, all LIBS spectra for the tea leaves have been collected and compared. The result was illustrated in Fig.3, and it was found that the profiles of the 6 kinds of tea leaves' LIBS spectra were in accordance with each other, which indicated that they have the same element compositions. But from the differences of the intensities of the characteristic spectra among the tea leaves, it can easily get the information that the content of the same element varies in different kinds of tea leaves. Choosing the appropriate characteristic spectra is very important for the classification results. In this work, several elements' characteristic spectra with significant differences were selected as the analysis indexes to classify the tea leaves. Through the comparative analysis of the tea leaves' LIBS spectra, the spectral intensities of Mg (279.55nm), Mn (279.83nm), CN (0-0) (388.34nm), Ca (393.37nm), Al (396.15nm), C_2 (0-0) (516.45nm), Fe (517.46nm) and K (766.49nm) were selected as the analytical lines. In addition, the intensity of C I 247.86 nm spectral line was selected as the reference in order to reduce the experimental error and improve the reliability of the results. The relative standard deviations (RSD) vary for the six kind of tea leaves, but the maximum RSD is less than 20%.

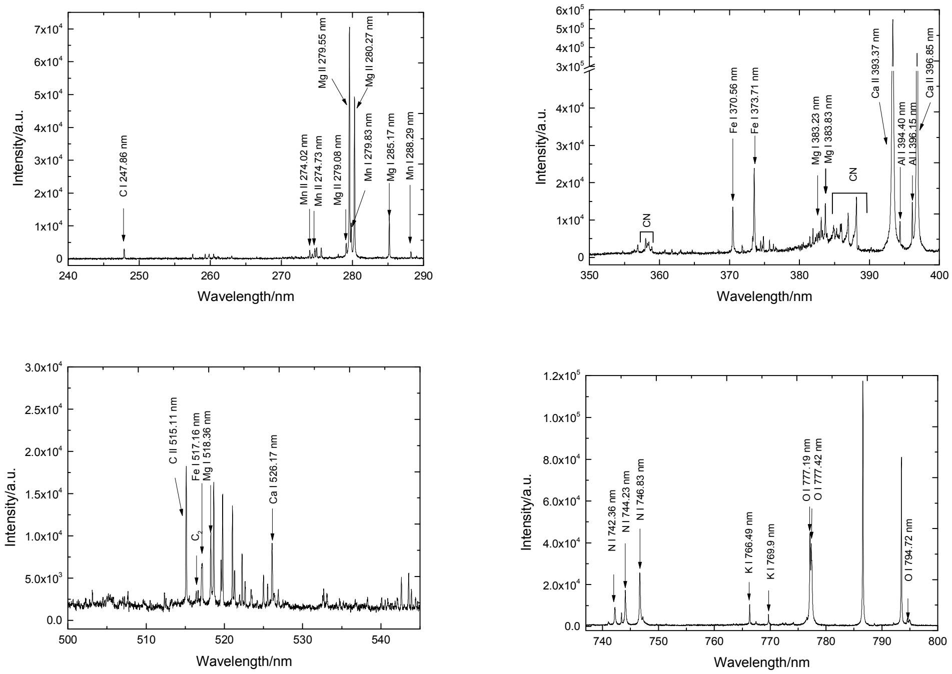

Fig.2 LIBS spectra of Longjing Green tea leaves in the wavelength range of 240~800 nm

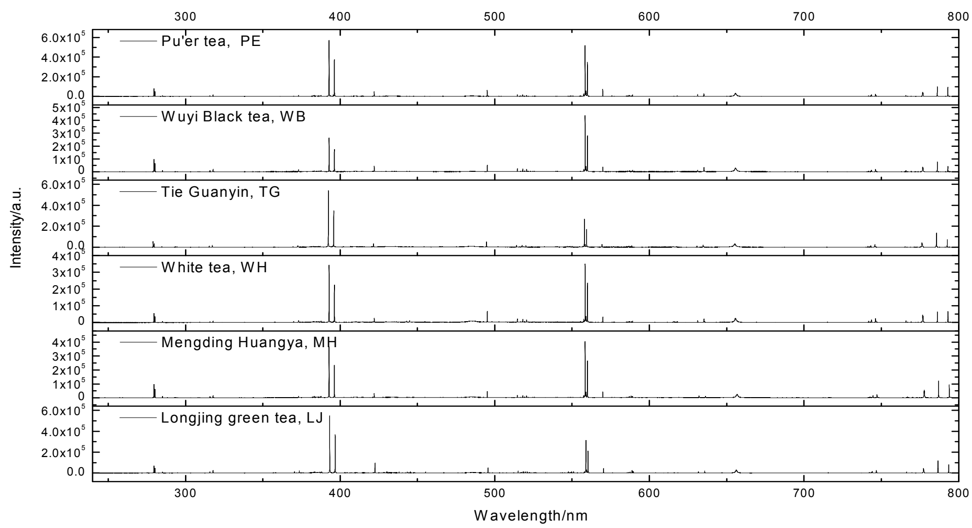

Fig.3 LIBS spectra of six kinds of Chinese tea in the open air

In the experiments, 100 spectra were recorded by spectrometer system for each kind of tea leaf, and 600 spectra were recorded in total. In order to find an operative classification rule for discriminating tea leaves, supervised learning pattern recognition techniques must be applied, such as stepwise LDA. This assumes a prior knowledge of the number of classes as well as the class membership of each sample in a training set. The variables included in the analysis are determined with a stepwise LDA, using a Wilk's lambda selection criterion and an F-statistic factor, to establish the significance of changes in lambda when a new variable is tested. Half of each tea leaf's LIBS spectra were used as training set to establish discriminant functions, and the remaining part was used as testing set to verify the accuracy of the results.

The results of the classification are output in chart form when using the SPSS statistics software package, and the classification results of the training set and testing set are shown in Fig 4. Figure 4 is the scatter diagram of all tea leaves, representing the scores of the first and second typical discriminant functions on each sample. The horizontal axis represents the score of the first typical variable, and the vertical axis is the score of second typical variable, and the fractional variables are calculated according to the sample data. The classification results show that the distribution of each class concentrated in a particular region, and scores for different classes vary. Class 1 (Longjing green tea) and class 6 (Pu'er tea) can obviously distinguish from class 2, 3, 4 and 5. Although there is some overlap in class 2(Mengding huangya tea), class 3(White tea) and class 4(Tie Guanyin) and class 5(Wuyi black tea), but it is still effective to classify the classes of tea leaves with high accuracy.

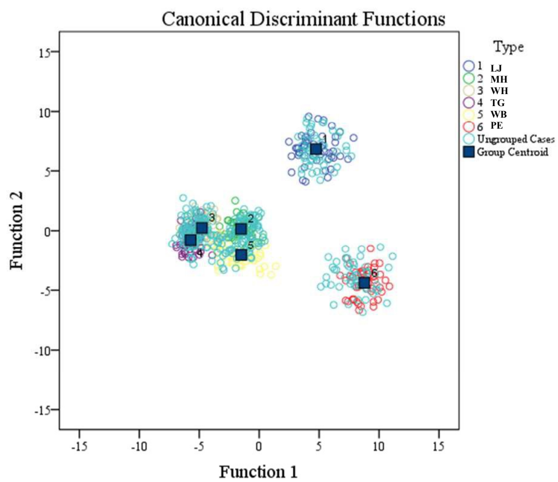

Fig.4 The scatter diagram contains all classes

The discriminat analysis results of the tea leaves are shown in detailed in table 1 and table 2, when discriminate the classes of the samples under the assistance of SPSS statistics software package, and it can identify the test samples and training samples as well. As to the returned classifications for training set in table 1, the correct identification rates for Longjing green tea, White tea, Tie guanyin and Pu'er tea are 100%, while the correct identification rates for Mengding huangya and Wuyi black tea are 96% and 92% respectively. Totally, 294 samples are identified correctly from 300 training set, and the average correct identification rate for returned classification reach up to 98%.

Table 2 shows the classification results for the testing set. The correct identification rates for Longjing green tea, Tie guanyin and Pu'er tea are 100%, while the correct identification rates for Mengding huangya White tea, and Wuyi black tea are 94%, 98% and 80% respectively. In addition, 286 samples from 300 test set are correctly identified, and the average correct identification rate is 95.33%.

The misclassification mainly occurred among Mengding huangya tea leaf, White tea leaf and Black tea leaf due to little difference of the line intensities of Mg, Mn, Ca, Al and C_2 (0-0) among these tea leaves. Overall, LIBS technique combined with discriminant analysis method has the advantages of fast, simplicity and accuracy, these research results provide available scheme for the detection and discrimination of agricultural products.

Table 1 Classification results of all samples [a, b]

	Type	Predicted group membership/%						Total
		LJ	MH	WH	TG	WB	PE	
Original	LJ	100.0	0.0	0.0	0.0	0.0	0.0	100.0
	MH	0.0	96.0	2.0	0.0	2.0	0.0	100.0
	WH	0.0	0.0	100.0	0.0	0.0	0.0	100.0
	TG	0.0	0.0	0.0	100.0	0.0	0.0	100.0
	WB	0.0	6.0	2.0	0.0	92.0	0.0	100.0
	PE	0.0	0.0	0.0	0.0	0.0	100.0	100.0
Cross-validated [c]	LJ	100.0	0.0	0.0	0.0	0.0	0.0	100.0
	MH	0.0	96.0	2.0	0.0	2.0	0.0	100.0
	WH	0.0	0.0	100.0	0.0	0.0	0.0	100.0
	TG	0.0	0.0	0.0	100.0	0.0	0.0	100.0
	WB	0.0	6.0	2.0	0.0	92.0	0.0	100.0
	PE	0.0	0.0	0.0	0.0	0.0	100.0	100.0

a. 98.0% of original grouped cases correctly classified.

b. 98.0% of cross-validated grouped cases correctly classified.

c. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

Table 2 Classification results for the testing samples

Type	Predicted group membership/%						Total
	LJ	MH	WH	TG	WB	PE	
LJ	100.0	0.0	0.0	0.0	0.0	0.0	100.0
MH	0.0	94.0	0.0	0.0	6.0	0.0	100.0
WH	0.0	2.0	98.0	0.0	0.0	0.0	100.0
TG	0.0	0.0	0.0	100.0	0.0	0.0	100.0
WB	0.0	16.0	4.0	0.0	80.0	0.0	100.0
PE	0.0	0.0	0.0	0.0	0.0	100.0	100.0

4. Conclusions

In this work, LIBS technique was employed to detect Chinese tea leaves, and LIBS combined with discriminant analysis method was performed to identify six kinds of tea leaves. Based on the preliminary analysis of tea leaves components, the fingerprint information of these six kinds of Chinese tea leaves was extracted. The spectral lines of Mg, Mn, Ca, Al, Fe, K, CN and C_2 were selected as the object of evaluation. The results show that 294 samples are identified correctly from 300 training set data, and the average correct identification rate is 98%. In addition, 286 samples from 300 test set data are correctly identified, and the average correct identification rate is 95.33%. Overall, LIBS technique combined with discriminant analysis method has the advantages of fast, simplicity and accuracy, and provide available scheme for the detection and discrimination of agricultural products.

Acknowledgements

This work is finically supported by the National Nature Science Foundation of China (61205149), Scientific and Technological Talents Training Project of Chongqing (CSTC2013kjrc-qnrc40002), Scientific Research Foundation for the Returned Overseas Chinese

Scholars, State Education Ministry, Visiting Scholarship of State Key Laboratory of Power Transmission Equipment & System Security and New Technology(Chongqing University)(2007DA10512714409) and the Key Project of Foundation and Advanced Technology Research Project of CQ CSTC (cstc2015jcyjB0358).

References

1 L. H. Li, Q. L. Fu, V. Achal, Y.L. Liu, *Environ. Monit. Assess.*, 2015, 187(5): 228

2 Chinese Standard, GB/T 30766-2014

3 B Paz-Rodríguez, M. R. Domínguez-González and M Aboal-Somoza, *Food Chem.*, 2015, **170,** 492-500.

4 Q. Han, S. Mihara and K. Hashimoto, *Food Sci. Tech. Research,* 2014, **20,** 1109-1119.

5 P. Ning, C. Gong and Y. Zhang, *Spectrosc. Spec. Anal.*, 2010, **30,** 2830-2833.

6 A. Kowalsick, N. Kfoury andA. Robbat, *J. Chromatogr. A*, 2014, **1370,** 230-239.

7 S. W. Kim, A. M. AEl-Aty and Y. J. Lee, *Food Sci. Biotech.*, 2014, **23,** 2097-2101.

8 L. F. Cai, C. X. Xu and M. H. Zhong, *Asian J. Chem.*, **25,**8169-8172.

9 Y. Chen, J. Deng and Y. Wang, *J. Sci. Food Agr.*, 2014, **94,** 1026-1033.

10 S. Eschlböck-Fuchs, M. J. Haslinger and A. Hinterreiter, *Spectrochim. Acta Part B*, 2013, **87,** 36-42.

11 P. Zheng, H. Liu, J. Wang and M. Shi, *J. Anal. Atom. Spectrom.*, 2015, **30,** 867-874.

12 S. Moncayo, S. Manzoor and T. Ugidos, *Spectrochim. Acta Part B*, 2014, **101,** 21-25.

13 L. Liang, T. Zhang and K. Wang, *Appl. Optics*, 2014, **53,**544-552.

14 V. K. Unnikrishnan, K. S. Choudhari and S. D. Kulkarni, *RSC Advances*, 2013, **3,** 25872-25880.

15 Y. Yang, Q. H. Zhong and L. Chang-Mao, *Acta Phys. Sin*, 2013, **62,** 215201(1)-215201(7).

16 P. Zheng, H. Liu, J. Wang, B. Yu, B. Zhang, R. Yang, and X. Wang, *Anal. Methods*, 2014, **6,** 2163-2169.

17 P. Zheng, S. Jie, J. Wang, and H. Liu, *Plasma Sci. Tech*. 2015, **17**, 664-670.

18 N. Mahmoudi and E. Duman, *Expert Syst. Appl.*, 2015, **42,** 2510-2516.

19 http://physics.nist.gov/PhysRefData/ASD/lines_form.html.

20 S. Morel, N. Leone, P. Adam, et al. *Applied Optics*, 2003, **42**: 6184-6191