

UC Davis

UC Davis Previously Published Works

Title

Quantification of toxic metals using machine learning techniques and spark emission spectroscopy

Permalink

<https://escholarship.org/uc/item/4429n1kd>

Journal

Atmospheric Measurement Techniques, 13(10)

ISSN

1867-1381

Authors

Davari, Seyyed Ali

Wexler, Anthony S

Publication Date

2020

DOI

10.5194/amt-13-5369-2020

Peer reviewed



Published in final edited form as:

Atmos Meas Tech. 2020 ; 13(10): 5369–5377. doi:10.5194/amt-13-5369-2020.

Quantification of toxic metals using machine learning techniques and spark emission spectroscopy

Seyyed Ali Davari¹, Anthony S. Wexler^{1,2,3}

¹Air Quality Research Center (AQRC), University of California, Davis, 95616, Davis, USA

²Department of Mechanical and Aerospace Engineering, Civil and Environmental Engineering, University of California, Davis, 95616, Davis, USA

³Land, Air and Water Resources, University of California, Davis, 95616, Davis, USA

Abstract

The United States Environmental Protection Agency (US EPA) list of hazardous air pollutants (HAPs) includes toxic metal suspected or associated with development of cancer. Traditional techniques for detecting and quantifying toxic metals in the atmosphere are either not real time, hindering identification of sources, or limited by instrument costs. Spark emission spectroscopy is a promising and cost-effective technique that can be used for analyzing toxic metals in real time. Here, we have developed a cost-effective spark emission spectroscopy system to quantify the concentration of toxic metals targeted by the US EPA. Specifically, Cr, Cu, Ni, and Pb solutions were diluted and deposited on the ground electrode of the spark emission system. The least absolute shrinkage and selection operator (LASSO) was optimized and employed to detect useful features from the spark-generated plasma emissions. The optimized model was able to detect atomic emission lines along with other features to build a regression model that predicts the concentration of toxic metals from the observed spectra. The limits of detections (LODs) were estimated using the detected features and compared to the traditional single-feature approach. LASSO is capable of detecting highly sensitive features in the input spectrum; however, for some toxic metals the single-feature LOD marginally outperforms LASSO LOD. The combination of low-cost instruments with advanced machine learning techniques for data analysis could pave the path forward for data-driven solutions to costly measurements.

1 Introduction

The United States Environmental Protection Agency (US EPA) lists a number of metals in their list of hazardous air pollutants (HAPs). These metals are known or suspected to cause cancer or other serious health effects (Buzzea et al., 2007; Pope et al., 2002). Table 1 lists the metals in the US EPA's HAP list. Table 2 lists other metals that are not on the US EPA's HAP list but have been implicated in a range of adverse health effects and so are of concern

This work is distributed under the Creative Commons Attribution 4.0 License.

Correspondence: Seyyed Ali Davari (seyyed.davari@arb.ca.gov).

Author contributions. ASW conceived the idea and advised SAD on this work. SAD carried out the experiments. SAD also conceived of the machine learning approach to analyzing the spectra.

Competing interests. The authors declare that they have no conflict of interest.

to the California Air Resources Board (CARB). It has been shown that presence of these metals is associated with various health concerns such as diabetes (Zanobetti et al., 2009), cardiovascular disease (Brook et al., 2004), and asthma (Gent et al., 2009). Therefore, it is necessary to monitor and quantify their ambient concentration.

Various techniques over the years have been developed and used to measure metal particles. X-ray fluorescence (XRF) (Vincze et al., 2002; Van Meel et al., 2007) and inductively coupled plasma mass spectrometry (ICP-MS) (Rovelli et al., 2018; Venecek et al., 2016) have been used traditionally to quantify metals in atmospheric particles. XRF is excellent for measuring lighter elements and metals on filter substrates, but for field application it is expensive, has a high limit of detection (LOD) for heavier elements, and includes radiation risk. ICP-MS requires collection of aerosol on a substrate, such as a filter or impactor foil, extraction of the metals or elements from the substrate using harsh acidic chemicals, and then analyzing in the ICP-MS along with standards that help the instrument quantitate. Moreover, ICP-MS is most suitable for heavier elements and metals so has a high LOD for lighter toxic metals and is not available in field-deployed, real-time applications. Additionally, these instruments are expensive and hence are limited by cost and complexity as well.

Spark-induced breakdown spectroscopy (SIBS) and laser-induced breakdown spectroscopy (LIBS) have been employed in various applications from combustion (Kotzagianni et al., 2016; Do and Carter, 2013; Kiefer et al., 2012), nanomaterials (De Giacomo et al., 2011; Hu et al., 2017; Davari et al., 2017a; Matsumoto et al., 2015a, b, 2016), and environmental and bio-hazards (Diwakar and Kulkarni, 2012; Diwakar et al., 2012; Zheng et al., 2018b), forensics (Martin et al., 2007), semiconductors and thin films (Hermann et al., 2019; Davari et al., 2017b; Axente et al., 2014; Davari et al., 2019), explosives (Gottfried et al., 2009), pharmaceuticals (Mukherjee and Cheng, 2008a, b; St-Onge et al., 2002), and biomedical engineering (Abbasi et al., 2018; Baudelet et al., 2006; Davari et al., 2018). In particular, Fisher et al. (2001) studied various toxic metals in aerosols by optimizing the spectrometer response with respect to gate delay. Hunter et al. (2000) employed spark emission spectroscopy for continuous monitoring of metallic elements in aerosols. Yao et al. (2018) used spark emission spectroscopy to obtain the carbon content of fly ashes. Diwakar and Kulkarni (2012) employed spark emission spectroscopy coupled with a corona aerosol microconcentrator (CAM) to improve the particle collection efficiency and detection limits of toxic metals. Zheng et al. (2017) characterized the CAM performance with respect to different experimental parameters and obtained the optimized design parameters for their CAM system.

Recently, machine learning and deep learning techniques have been applied in different fields. These techniques in general learn patterns that can be used to distinguish different labels. Boucher et al. (2015) employed various linear and nonlinear machine learning techniques on LIBS spectra obtained from geological samples and concluded that a combination of models yields a lower total error of prediction. Chengxu et al. (2018) used convolutional neural networks to detect potassium in LIBS spectra and improve the linearity of their prediction model incorporating deep convolutional layers. Zheng et al. (2018a) employed spark emission spectroscopy on metals and used partial least squares regression to

analyze their spectra set. They compared their multivariate models to univariate models and showed in their study that these two groups have similar performance.

While LIBS and SIBS address issues regarding the field measurement and instrument complexity, they are still considered expensive. Current interest in low-cost sensors and their ability to characterize local air pollution concentrations motivated development of a low-cost system. We employed two complementary approaches: (1) decreasing the cost of the electronics associated with SIBS and (2) incorporating advanced data analysis techniques to improve quantification and the limit of detection. In recent years, numerous studies have used artificial neural networks (Ferreira et al., 2008), partial least squares regression, and least absolute shrinkage and selection operator (LASSO) (Dyar et al., 2012) on emission spectra to improve the quantification and limit of the detection of spectroscopic systems. In this study, we have developed a low-cost spark emission spectroscopy system to quantify toxic metals. To reduce the overall cost, inexpensive replacements for necessary components, such as the spark generator and delay generator, have been developed in the lab. To improve performance, advanced machine learning tools such as k -means clustering and LASSO have been employed to improve the system performance. The resulting instrument was evaluated against four toxic metals listed by the US EPA.

2 Instrument development

2.1 Spark generation system

Setting up a spark emission spectroscopy system requires expensive components. However, depending on the application, some of the components can be replaced. Components such as spark generator and delay generator can cost up to USD 10 000 and 5000, respectively. According to our application and needs, we developed these components for less than USD 600 and 50, respectively. One costly component that is required for developing a spark emission spectroscopy system is the spark generation system. Numerous papers have studied the fundamental principles of spark emission spectroscopy (Walters, 1977; Sacks and Walters, 1970; Walters, 1969). The key idea is to discharge a capacitor as quickly as possible to increase the power dissipated in the spark gap. Figure 1 illustrates the schematic of the spark generation system. The overall goal is to charge a capacitor at high voltage and, once it has been charged sufficiently, discharge the capacitor through the spark gap. An Arduino board controls the timing between charging and discharging the capacitor. A boost convertor converts 24 to 5000 V DC and is connected to a mechanical relay with two switching states controlled with the Arduino board. In the charge state, the mechanical relay provides the conduction path between the boost convertor and the capacitor. In this configuration, the capacitor reaches full charge in 5 s. Once the capacitor is fully charged, the Arduino board sends a signal to turn off the boost convertor and sends another signal to the mechanical relay to flip to the discharge state. At the discharge state, the mechanical relay provides a conduction path between the capacitor and the spark gap. Shepherd et al. (2000) showed that the discharge process could be controlled by a resistor after the spark gap. For low resistor values, the spark current exhibited a periodic behavior as the capacitor discharges, which can be associated with an under-damped discharging. On the other hand, increasing the resistor value damped the discharge process and dissipated a large portion

of the capacitor energy through the resistor instead of the spark gap. We found that a 10 Ω resistor maximizes the power dissipation in the spark gap, while minimizing oscillations. Figure 2 illustrates the evolution of the generated spark as a function of time. The voltage shows a sudden increase followed by an exponential decrease fully discharging in less than 5 μs and thus delivering sufficient energy to the arc and deposited analyte.

2.2 Delay generator

The delay generator is another costly component typically used in time-resolved spectroscopy. Electronics advances have paved the way for developing a cost-effective delay generator. The delay generator suppresses initial noise in the emission spectrum so needs to cover a range between 1 and 20 μs with resolution less than 0.2 μs . We designed a custom-built delay generator in order to lower the overall cost of the instrument. Figure 3 illustrates the schematic of the circuit. Upon generation of the spark-induced plasma, a pair of lenses collects and focuses the plasma emission into a photodiode. The pulse generated by the photodiode is passed into a voltage comparator (LM 311-N) to generate a transistor–transistor logic (TTL) signal. The output TTL signal from the comparator is sent to a pulse width modulator (PWM) controller (LTC6992), which adds delay to the TTL signal. An Arduino board adjusts a digital resistor (AD5241), which in turn determines the delay value. Figure 4 shows the delay generator performance. The y axis illustrates the delay values requested of the delay generator, while the x axis shows the measured values. The red dashed line shows the desired 1 : 1 line, while the circles show the measured performance. The performance is linear over the relevant delay range with only a slight deviation from the 1 : 1 line. Considering the spark-generated plasma's short lifetime, our measurements require short delay values ($< 5 \mu\text{s}$) where the built-in delay generator shows excellent performance and accuracy.

2.3 Spectra collection

Four toxic metals with different concentrations were used to test the developed spark emission spectrometer system performance. Cr, Cu, Ni, and Pd ($1000 \mu\text{g mL}^{-1}$) were purchased from AccuStandard and diluted to specific concentrations. For each concentration more than 10 spectra have been collected and used for model development. A micropipette was used to deposit diluted solutions on a 1 mm diameter Tungsten ground electrode of the spark system for emission analysis. The total mass can be calculated from the deposited volume and solution concentration. Upon evaporation of the droplets, the capacitor was discharged to ablate the deposited material and obtain spectra. A pair of lenses (75 mm focal length and 25.4 mm diameter; Thorlab) focused the emission into an optical fiber connected to a spectrometer (Ocean Optics).

3 Results and discussions

To address shot-to-shot variations in the spark-generated plasma and nullify possible faults caused by the low-cost components, an unsupervised learning technique, k -means clustering, classifies the collected spectra. Following this procedure, it is possible to identify and remove outliers and hence improve the accuracy of the analysis. Figure 5 illustrates the elbow plot that is used to optimize the number of spectral classes. The standard approach is

to set the optimum number of clusters to the value where the within-cluster sum of squares (WCSS) error plateaus. The WCSS error plateaus once we have two or more centroids, and, therefore, the number of centroids is set to two. Figure 6 illustrates the performance of the model for 300 spectra obtained from the background (Tungsten ground electrode ablation). The results show clearly two clusters with different emission response. The lower left cluster containing < 10 % of the spectra represents low-signal outliers, which were eliminated from further analysis. For each toxic metal, 0.1, 1, 10, and 100 ng of mass were deposited on the ground electrode. For each concentration, 10 spectra were collected using a 2 μ s delay between the observed and recorded emissions. After ablating the deposited mass and recording the spectrum, feature scaling has been used as a preprocessing step to improve the optimization process for our machine learning model. Plasma temperature can be obtained as follows:

$$I_{em} = \frac{hc}{\lambda_{ki}} N_k A_{ki}, \quad (1)$$

$$N_k = N \frac{g_k e^{-\frac{E_k}{k_B T}}}{U(T)}. \quad (2)$$

Combining Eqs. (1) and (2) and taking log from both sides gives

$$\ln\left(\frac{I_{em} \lambda_{ki}}{g_k A_{ki}}\right) = -\frac{E_k}{k_B T} + \ln\left(\frac{hcN}{U(T)}\right), \quad (3)$$

where k_B is the Boltzmann constant, A_{ki} is the transition probability between two energy states (i) and (k), and N_k is the population density at energy state k (E_k). λ_{ki} indicates the wavelength associated with the transition, and g_k represents the degeneracy of energy state k . The slope of Eq. (3) is used to estimate the plasma temperature based on a series of Tungsten lines for the recorded cleaned spectra set at 2 μ s. Figure 7 illustrates the Boltzmann plot (Hahn and Omenetto, 2010, 2012) constructed by Tungsten lines. Based on the slope of the fit, the plasma temperature is estimated as 4013 \pm 579 K. Upon identifying and removing the outlier spectra, the cleaned spectra set is normalized using the Tungsten peak at W I (400.87 nm) and fed into the LASSO algorithm for model development and prediction.

LASSO

The cleaned scaled spectra set has been used to detect and quantify concentrations of the toxic metals. Simple linear regression obtains the slope and intercept of a linear line by minimizing the mean squared error between the predictions and known values. LASSO detects and employs more features to perform predictions by optimizing the following loss function:

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m (y^{(i)} - h_{\theta}(\mathbf{x}^{(i)}))^2 + c \sum_{j=1}^k |\theta_j|, \quad (4)$$

where $\mathbf{x}^{(i)} \in \mathbb{R}^{2048}$ and $h_{\theta}(\mathbf{x}^{(i)})$ represent the normalized spectrum and the LASSO concentration prediction based on spectrum (i) ($\mathbf{x}^{(i)}$), respectively, and where $y^{(i)}$ is the known concentration corresponding to spectrum (i). m refers to the total number of spectra, and the LASSO coefficients are indicated by θ_j . k indicates the total number of features (spectral lines) used to build the model. The first term in Eq. (4) is the mean squared error and is common with simple linear regression, while the second term is a regularization term that minimizes the magnitude of θ_j . The L1 norm essentially sets most of the features in the spectrum to zero and maintains only a few features to build the linear model and perform predictions. The regularization constant (c) determines the number of features to be used in the model, and therefore the model loss needs to be optimized with respect to the regularization constant. To obtain the optimized regularization constant, we plotted the loss values for the Ni spectra training and testing sets as a function of the number of features for various c values based on leave-one-out cross validation (Fig. 8). As expected, the train loss monotonically decreases as the number of features increases, while the loss for the test set initially decreases and then starts increasing. This implies that after incorporating a certain number of features into the model, the model starts memorizing rather than generalizing, which is known as overfitting. Therefore, we set the regularization constant to the value that minimizes the loss for the test set. Figure 9 illustrates the optimized LASSO model predictions obtained by cross validation. For each concentration, the cross validation predictions were averaged and plotted along with the standard deviations. The predicted values vary linearly with the actuals. Figure 10 shows the wavelengths chosen by LASSO and the mean spectrum for 10 ng. LASSO chose a few Ni emission peaks along with other features to build the model. The same optimization process was applied to other toxic metals, specifically Cr, Cu, and Pb. Figure 11 illustrates the resulting predictions and demonstrates the value of LASSO for predicting deposited mass from the spectra. To obtain the LOD, the following function of the LASSO coefficients θ_j was used:

$$\text{LOD} = 3 \frac{\sigma_B}{S} = 3 \sigma_B \|\theta_B\|, \quad (5)$$

where σ_B is the standard deviation of the background and $\|\theta_B\|$ is the Euclidean norm of LASSO coefficients. Table 3 reports the LODs of the studied toxic metals.

Multivariate regression models such as LASSO might be more powerful in detection and quantification over univariate models; however, there is no guarantee that multivariate models outperform simple linear regression (Castro and Pereira-Filho, 2016; Braga et al., 2010). To evaluate LASSO performance, we compared LASSO with univariate methods, by calculating the LODs using simple univariate linear regression based on the features selected by LASSO. Figure 12 illustrates the LODs obtained using this univariate technique (circles)

compared to LASSO LOD (dashed line) for Ni. Considering only the sensitivity (LOD) is necessary but not sufficient for evaluating model performance since low R^2 values are also problematic. Therefore, in order to incorporate both R^2 and LOD for model assessment, we defined a score as

$$\text{score} = \left(\frac{\text{LOD}}{R^2} \right)^2. \quad (6)$$

Based on this definition, a model that has low LOD and high R^2 is desirable. The LASSO score outperforms single feature linear regression for Pb, but the two methods were comparable for Cu, Ni, and Cr (Fig. 13). Other studies have reported that univariate techniques performed better than multivariate ones (Castro and Pereira-Filho, 2016; Braga et al., 2010). In LASSO, this may be related to the cost function defined for the regression (Eq. 4). LASSO is a special case of elastic net family where both L1 and L2 norms are combined and used in the cost function. Considering the cost function in Eq. (4), the model goal is to minimize the prediction error and coefficient values (minimizing L1). This does not necessarily optimize LOD. Therefore, cost function minimization does not correspond to LOD minimization. Considering Fig. 12, using features defined by LASSO in a univariate model may yield better LOD than that obtained by LASSO alone. This might be an advantageous approach if the physical intuition of the features is not as important as the detection of toxic metals.

4 Conclusion

A cost-effective spark emission spectroscopy instrument was designed and developed to quantify toxic metals targeted by the US EPA and the California Air Resources Board. Costly components such as the spark generation system and delay generator were developed to lower the overall cost. An unsupervised learning technique was employed to detect outlier spectra. The cleaned spectra set was fed into LASSO for predicting the concentration of deposited samples on the ground electrode of the spark system from spectra obtained from the plasma. A combination of LASSO feature detection with univariate regression might improve the detection limits. Our results illustrate the promising realm of cost-effective sensors combined with advanced machine learning techniques to provide data-driven solutions to the traditional challenging problems.

Financial support.

This research has been supported by the California Air Resources Board (CARB).

Data availability.

Please contact the corresponding author for the data.

References

- Abbasi H, Rauter G, Guzman R, Cattin PC, and Zam A: Laser-induced breakdown spectroscopy as a potential tool for autocarbonization detection in laserosteotomy, *J. Biomed. Opt.*, 23, 071206, 10.1117/1.JBO.23.7.071206, 2018
- Axente E, Hermann J, Socol G, Mercadier L, Beldjilali SA, Cirisan M, Luculescu CR, Ristoscu C, Mihailescu IN, and Craciun V: Accurate analysis of indium–zinc oxide thin films via laser-induced breakdown spectroscopy based on plasma modeling, *J. Anal. Atom. Spectrom.*, 29, 553–564, 2014.
- Baudelet M, Guyon L, Yu J, Wolf J-P, Amodeo T, Fréjafon E, and Laloi P: Femtosecond time-resolved laser-induced breakdown spectroscopy for detection and identification of bacteria: A comparison to the nanosecond regime, *J. Appl. Phys.*, 99, 084701, 10.1063/1.2187107, 2006.
- Boucher TF, Ozanne MV, Carosino ML, Dyar MD, Mahadevan S, Breves EA, Lepore KH, and Clegg SM: A study of machine learning regression methods for major elemental analysis of rocks using laser-induced breakdown spectroscopy, *Spectrochim. Acta B*, 107, 1–10, 2015.
- Braga JWB, Trevizan LC, Nunes LC, Rufini IA, Santos D Jr, and Krug FJ: Comparison of univariate and multivariate calibration for the determination of micronutrients in pellets of plant materials by laser induced breakdown spectrometry, *Spectrochim. Acta B*, 65, 66–74, 2010.
- Brook RD, Franklin B, Cascio W, Hong Y, Howard G, Lipsett M, Luepker R, Mittleman M, Samet J, Smith SC Jr., and Tager I: Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association, *Circulation*, 109, 2655–2671, 2004. [PubMed: 15173049]
- Buza C, Pacheco II, and Robbie K: Nanomaterials and nanoparticles: sources and toxicity, *Biointerphases*, 2, MR17–MR71, 2007. [PubMed: 20419892]
- Castro JP and Pereira-Filho ER: Twelve different types of data normalization for the proposition of classification, univariate and multivariate regression models for the direct analyses of alloys by laser-induced breakdown spectroscopy (LIBS), *J. Anal. Atom. Spectrom.*, 31, 2005–2014, 2016.
- Chengxu L, Bo W, Jiang X, Zhang J, Kang N, and Yanwei Y: Detection of K in soil using time-resolved laser-induced breakdown spectroscopy based on convolutional neural networks, *Plasma Sci. Technol.*, 21, 034014, 10.1088/2058-6272/aaef6e, 2018.
- Davari SA, Hu S, and Mukherjee D: Calibration-free quantitative analysis of elemental ratios in intermetallic nanoalloys and nanocomposites using Laser Induced Breakdown Spectroscopy (LIBS), *Talanta*, 164, 330–340, 2017a. [PubMed: 28107937]
- Davari SA, Hu S, Pamu R, and Mukherjee D: Calibration-free quantitative analysis of thin-film oxide layers in semiconductors using laser induced breakdown spectroscopy (LIBS), *J. Anal. Atom. Spectrom.*, 32, 1378–1387, 2017b.
- Davari SA, Masjedi S, Ferdous Z, and Mukherjee D: In-vitro analysis of early calcification in aortic valvular interstitial cells using Laser-Induced Breakdown Spectroscopy (LIBS), *J. Biophotonics*, 11, e201600288, 10.1002/jbio.201600288, 2018.
- Davari SA, Taylor PA, Standley RW, and Mukherjee D: Detection of interstitial oxygen contents in Czochralski grown silicon crystals using internal calibration in laser-induced breakdown spectroscopy (LIBS), *Talanta*, 193, 192–198, 2019. [PubMed: 30368290]
- De Giacomo A, De Bonis A, Dell’Aglia M, De Pascale O, Gaudio R, Orlando S, Santagata A, Senesi G, Taccogna F, and Teghil R: Laser ablation of graphite in water in a range of pressure from 1 to 146 atm using single and double pulse techniques for the production of carbon nanostructures, *J. Phys. Chem. C*, 115, 5123–5130, 2011.
- Diwakar P, Kulkarni P, and Birch ME: New approach for near-real-time measurement of elemental composition of aerosol using laser-induced breakdown spectroscopy, *Aerosol Sci. Tech.*, 46, 316–332, 2012.
- Diwakar PK and Kulkarni P: Measurement of elemental concentration of aerosols using spark emission spectroscopy, *J. Anal. Atom. Spectrom.*, 27, 1101–1109, 2012.
- Do H and Carter C: Hydrocarbon fuel concentration measurement in reacting flows using short-gated emission spectra of laser induced plasma, *Combust. Flame*, 160, 601–609, 2013.

- Dyar M, Carmosino M, Breves E, Ozanne M, Clegg S, and Wiens R: Comparison of partial least squares and lasso regression techniques as applied to laser-induced breakdown spectroscopy of geological samples, *Spectrochim. Acta B*, 70, 51–67, 2012.
- Ferreira EC, Milori DM, Ferreira EJ, Da Silva RM, and Martin-Neto L: Artificial neural network for Cu quantitative determination in soil using a portable laser induced breakdown spectroscopy system, *Spectrochim. Acta B*, 63, 1216–1220, 2008.
- Fisher BT, Johnsen HA, Buckley SG, and Hahn DW: Temporal gating for the optimization of laser-induced breakdown spectroscopy detection and analysis of toxic metals, *Appl. Spectrosc*, 55, 1312–1319, 2001.
- Gent JF, Koutrakis P, Belanger K, Triche E, Holford TR, Bracken MB, and Leaderer BP: Symptoms and medication use in children with asthma and traffic-related sources of fine particle pollution, *Environ. Health Persp*, 117, 1168–1174, 2009.
- Gottfried JL, De Lucia FC, Munson CA, and Miziolek AW: Laser-induced breakdown spectroscopy for detection of explosives residues: a review of recent advances, challenges, and future prospects, *Anal. Bioanal. Chem*, 395, 283–300, 2009. [PubMed: 19418042]
- Hahn DW and Omenetto N: Laser-induced breakdown spectroscopy (LIBS), part I: review of basic diagnostics and plasma–particle interactions: still-challenging issues within the analytical plasma community, *Appl. Spectrosc*, 64, 335A–366A, 2010. [PubMed: 21144145]
- Hahn DW and Omenetto N: Laser-induced breakdown spectroscopy (LIBS), part II: review of instrumental and methodological approaches to material analysis and applications to different fields, *Appl. Spectrosc*, 66, 347–419, 2012. [PubMed: 22449322]
- Hermann J, Axente E, Pelascini F, and Craciun V: Analysis of Multi-elemental Thin Films via Calibration-Free Laser-Induced Breakdown Spectroscopy, *Anal. Chem*, 91, 2544–2550, 2019. [PubMed: 30615420]
- Hu S, Ribeiro EL, Davari SA, Tian M, Mukherjee D, and Khomami B: Hybrid nanocomposites of nanostructured Co_3O_4 interfaced with reduced/nitrogen-doped graphene oxides for selective improvements in electrocatalytic and/or supercapacitive properties, *Rsc Advances*, 7, 33166–33176, 2017.
- Hunter AJ, Morency JR, Senior CL, Davis SJ, and Fraser ME: Continuous emissions monitoring using spark-induced breakdown spectroscopy, *J. Air Waste Manage*, 50, 111–117, 2000.
- Kiefer J, Tröger JW, Li Z, Seeger T, Alden M, and Leipertz A: Laser-induced breakdown flame thermometry, *Combust. Flame*, 159, 3576–3582, 2012.
- Kotzagianni M, Yuan R, Mastorakos E, and Couris S: Laser-induced breakdown spectroscopy measurements of mean mixture fraction in turbulent methane flames with a novel calibration scheme, *Combust. Flame*, 167, 72–85, 2016.
- Martin MZ, Labbé N, André N, Harris R, Ebinger M, Wullschlegler SD, and Vass AA: High resolution applications of laser-induced breakdown spectroscopy for environmental and forensic applications, *Spectrochim. Acta B*, 62, 1426–1432, 2007.
- Matsumoto A, Tamura A, Honda T, Hirota T, Kobayashi K, Katakura S, Nishi N, Amano K. i., Fukami K, and Sakka T: Transfer of the species dissolved in a liquid into laser ablation plasma: an approach using emission spectroscopy, *J. Phys. Chem. C*, 119, 26506–26511, 2015a.
- Matsumoto A, Tamura A, Koda R, Fukami K, Ogata YH, Nishi N, Thornton B, and Sakka T: On-site quantitative elemental analysis of metal ions in aqueous solutions by underwater laser-induced breakdown spectroscopy combined with electrodeposition under controlled potential, *Anal. Chem*, 87, 1655–1661, 2015b. [PubMed: 25560224]
- Matsumoto A, Tamura A, Koda R, Fukami K, Ogata YH, Nishi N, Thornton B, and Sakka T: A calibration-free approach for on-site multi-element analysis of metal ions in aqueous solutions by electrodeposition-assisted underwater laser-induced breakdown spectroscopy, *Spectrochim. Acta B*, 118, 45–55, 2016.
- Mukherjee D and Cheng M-D: Characterization of carbon-containing aerosolized drugs using laser-induced breakdown spectroscopy, *Appl. Spectrosc*, 62, 554–562, 2008a. [PubMed: 18498697]
- Mukherjee D and Cheng M-D: Quantitative analysis of carbonaceous aerosols using laser-induced breakdown spectroscopy: a study on mass loading induced plasma matrix effects, *J. Anal. Atom. Spectrom*, 23, 119–128, 2008b.

- Pope CA III, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, and Thurston GD: Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution, *Jama*, 287, 1132–1141, 2002. [PubMed: 11879110]
- Rovelli S, Nischkauer W, Cavallo DM, and Limbeck A: Multi-element analysis of size-segregated fine and ultrafine particulate via Laser Ablation-Inductively Coupled Plasma-Mass Spectrometry, *Anal. Chim. Acta*, 1043, 11–19, 2018. [PubMed: 30392658]
- Sacks RD and Walters JP: Short-time, spatially-resolved radiation processes in a high-voltage spark discharge, *Anal. Chem*, 42, 61–84, 1970.
- Shepherd JE, Krok JC, and Lee JJ: Spark ignition energy measurements in Jet A, California Institute of Technology, available at: http://www.twa800.com/report/galcit_fm97-9.pdf, last access: 1 October 2020.
- St-Onge L, Kwong E, Sabsabi M, and Vadas E: Quantitative analysis of pharmaceutical products by laser-induced breakdown spectroscopy, *Spectrochim. Acta B*, 57, 1131–1140, 2002.
- Van Meel K, Smekens A, Behets M, Kazandjian P, and Van Grieken R: Determination of platinum, palladium, and rhodium in automotive catalysts using high-energy secondary target X-ray fluorescence spectrometry, *Anal. Chem*, 79, 6383–6389, 2007. [PubMed: 17628117]
- Venecek MA, Zhao Y, Mojica J, McDade CE, Green PG, Kleeman MJ, and Wexler AS: Characterization of the 8-stage Rotating Drum Impactor under low concentration conditions, *J. Aerosol Sci*, 100, 140–154, 2016.
- Vincze L, Somogyi A, Osan J, Vekemans B, Török S, Janssens K, and Adams F: Quantitative trace element analysis of individual fly ash particles by means of X-ray microfluorescence, *Anal. Chem*, 74, 1128–1135, 2002. [PubMed: 11924974]
- Walters J: Historical advances in spark emission spectroscopy, *Appl. Spectrosc*, 23, 317–331, 1969.
- Walters JP: Spark discharge: Application multielement spectro-chemical analysis, *Science*, 198, 787–797, 1977. [PubMed: 17843391]
- Yao S, Xu J, Zhang L, Zhao J, and Lu Z: Optimizing critical parameters for the directly measurement of particle flow with PFSIBS, *Sci. Rep*, 8, 1868, 10.1038/s41598-018-20073-1, 2018. [PubMed: 29382884]
- Zanobetti A, Franklin M, Koutrakis P, and Schwartz J: Fine particulate air pollution and its components in association with cause-specific emergency admissions, *Environ. Health*, 8, 58, 10.1186/1476-069X-8-58, 2009. [PubMed: 20025755]
- Zheng L, Kulkarni P, Zavvos K, Liang H, Birch ME, and Dionysiou DD: Characterization of an aerosol microconcentrator for analysis using microscale optical spectroscopies, *J. Aerosol Sci*, 104, 66–78, 2017. [PubMed: 28626243]
- Zheng L, Kulkarni P, and Dionysiou DD: Calibration approaches for the measurement of aerosol multielemental concentration using spark emission spectroscopy, *J. Anal. Atom. Spectrom*, 33, 404–412, 2018a.
- Zheng L, Kulkarni P, and Diwakar P: Spatial and temporal dynamics of a pulsed spark microplasma used for aerosol analysis, *Spectrochim. Acta B*, 144, 55–62, 2018b.

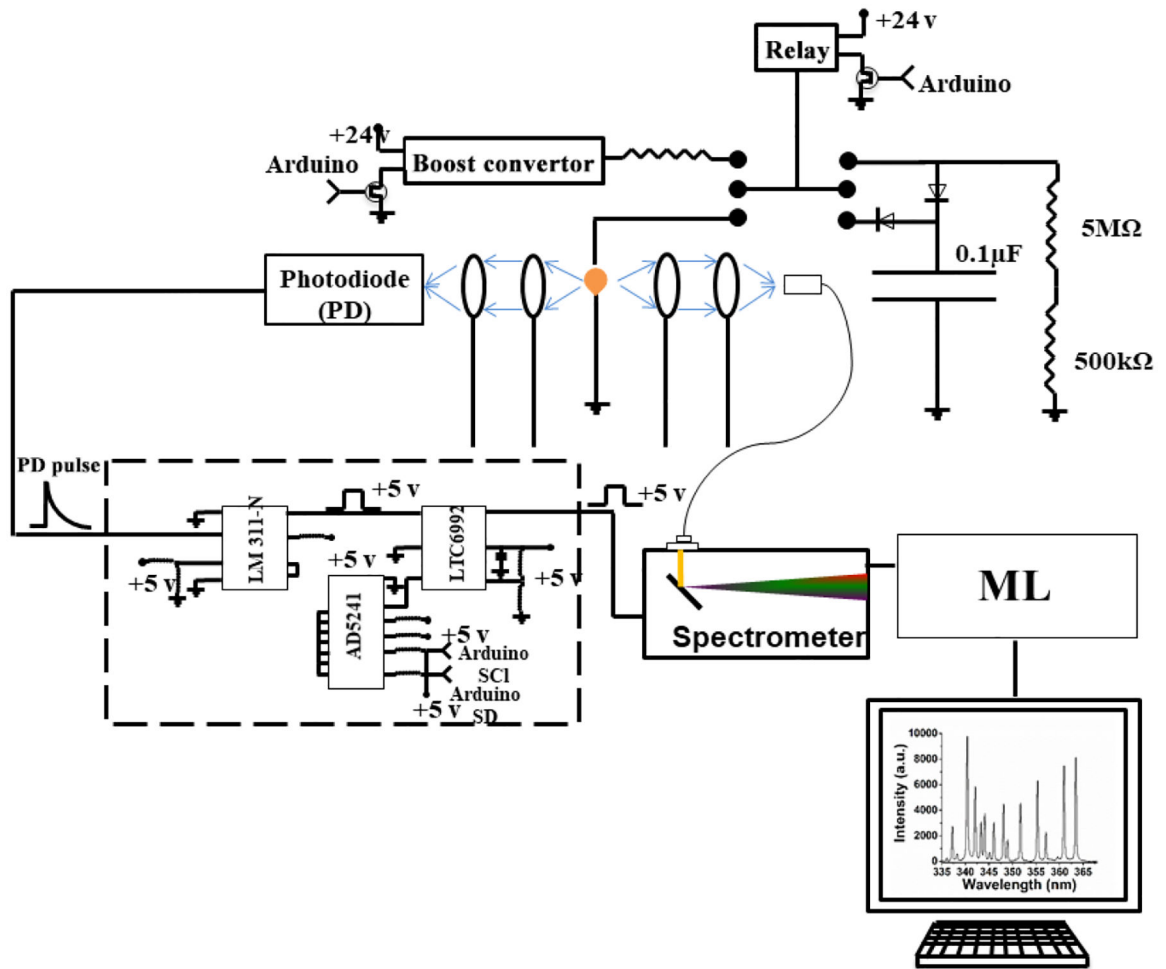


Figure 1. Schematic of the developed spark emission spectroscopy.

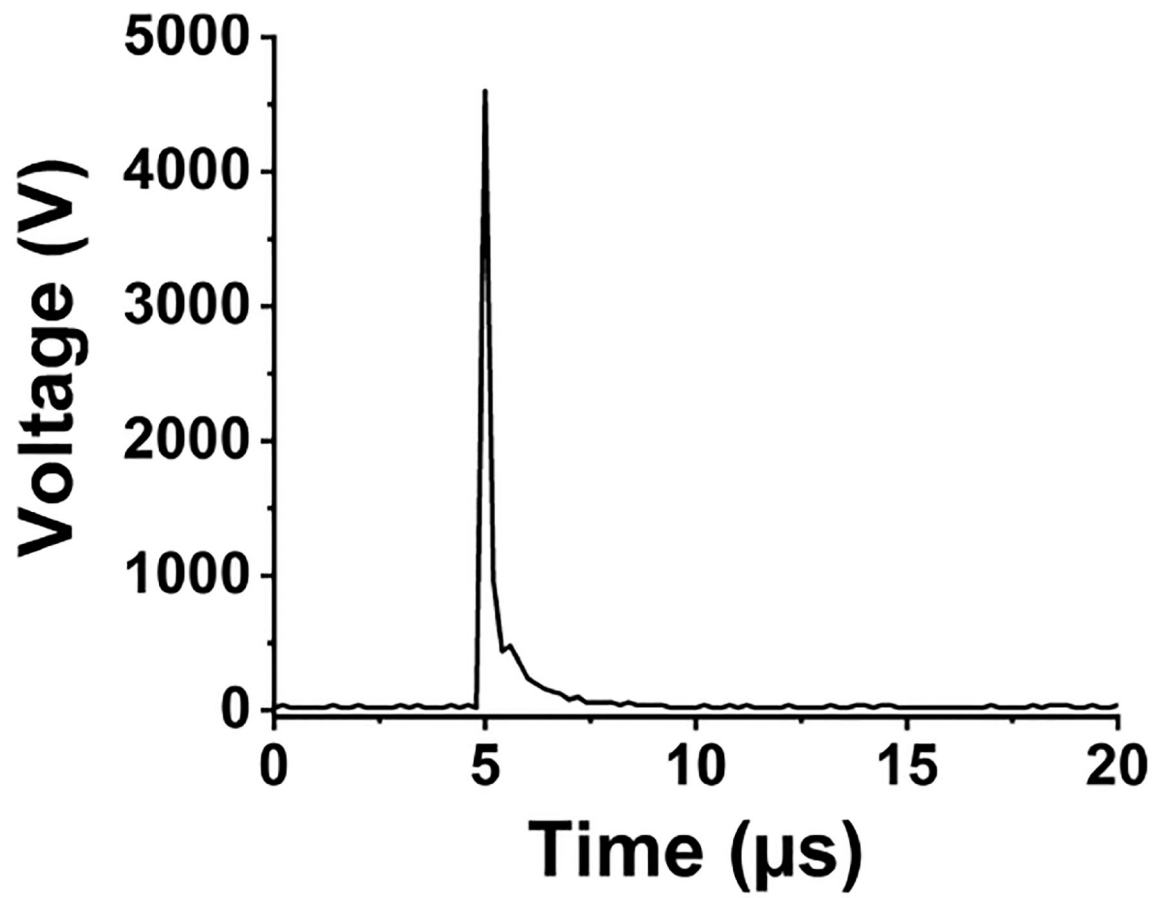


Figure 2.
Spark voltage evolution in time.

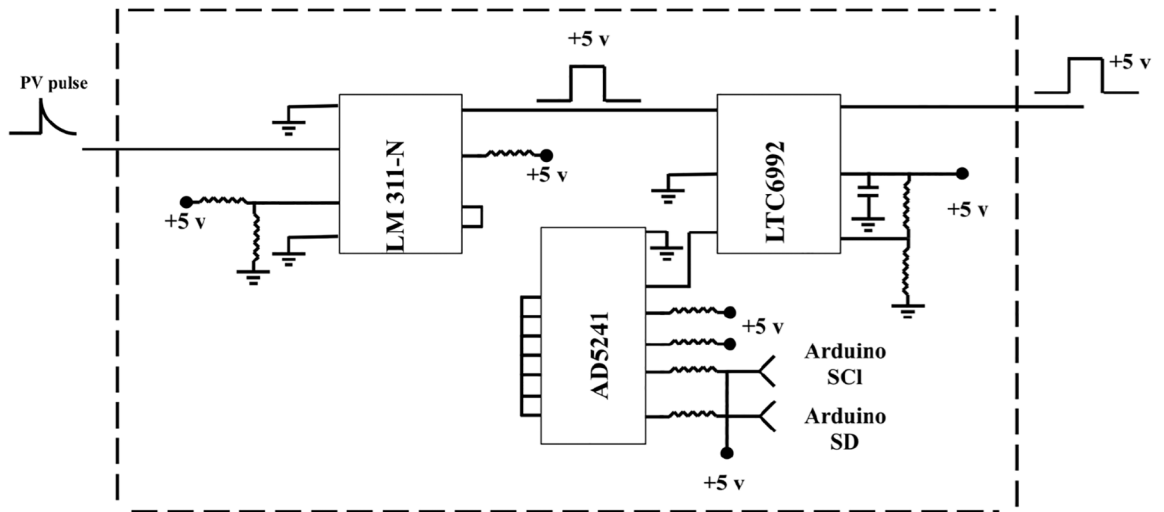


Figure 3.
Schematic of the built-in delay generator.

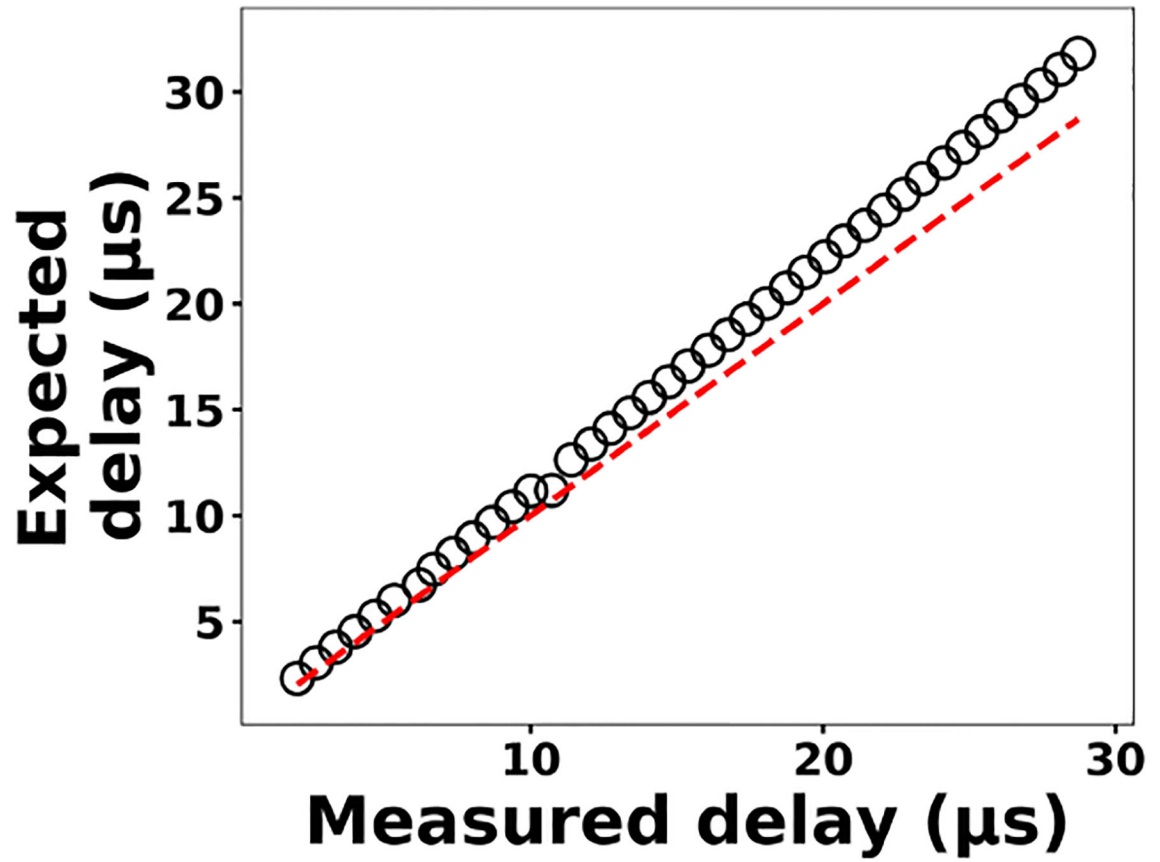


Figure 4.
The expected delay set by the Arduino board as a function of the measured delay.

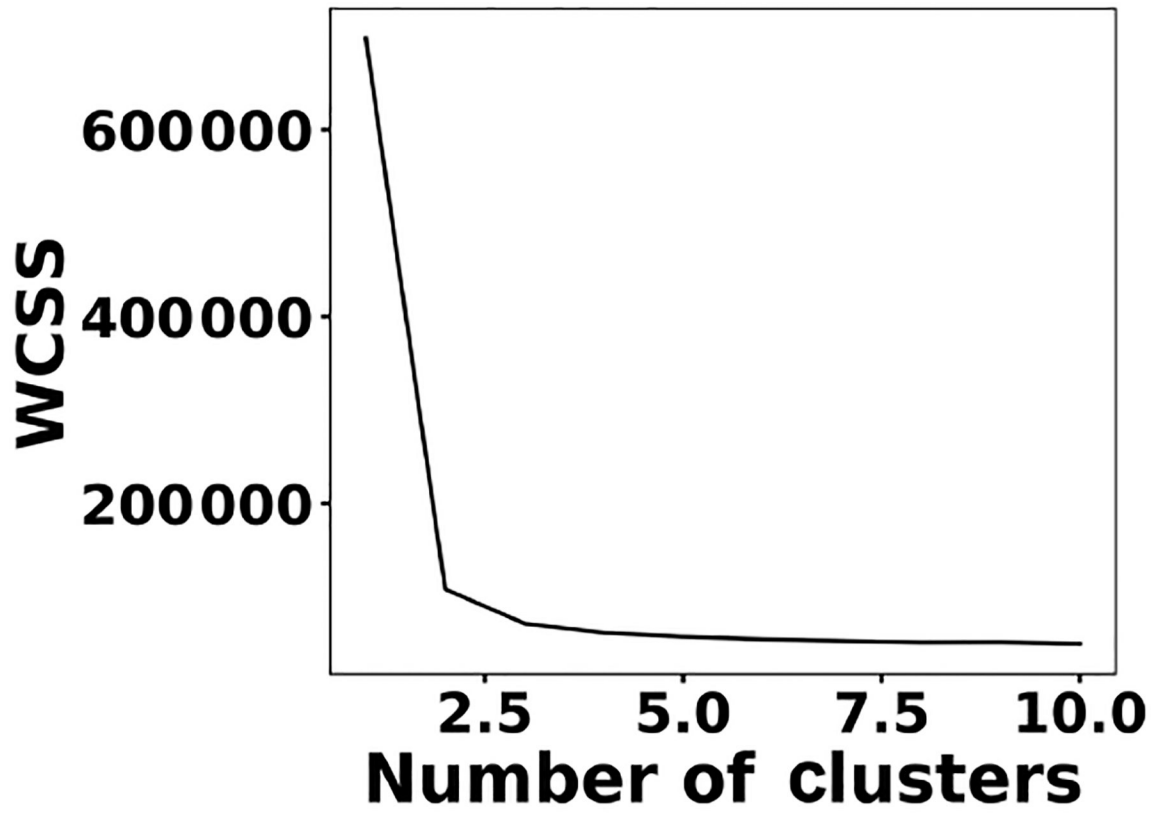


Figure 5.
The elbow plot suggests two centroids for clustering the spectra set.

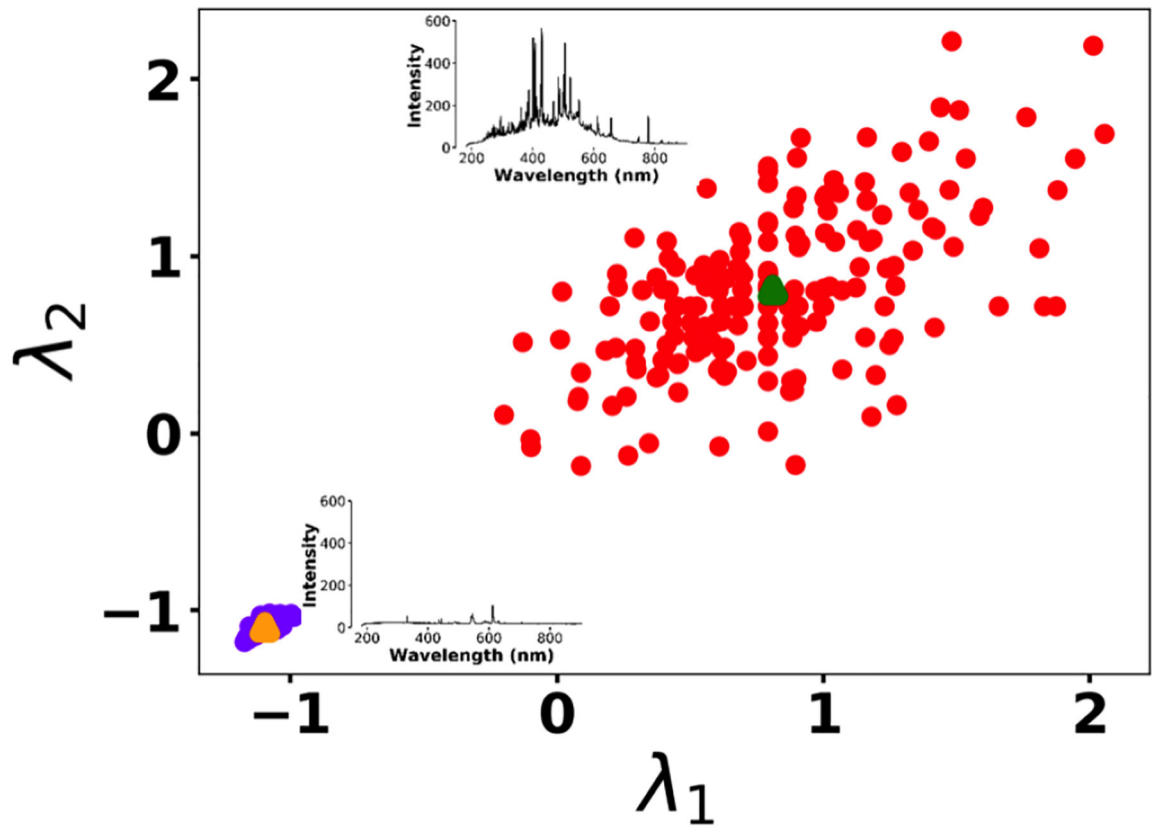


Figure 6.

K-means clustering for detecting outliers before passing the spectra set to the LASSO model. Two clusters were plotted for the normalized intensities of two arbitrary wavelengths at λ_1 (208.365 nm) and λ_2 (208.759 nm).

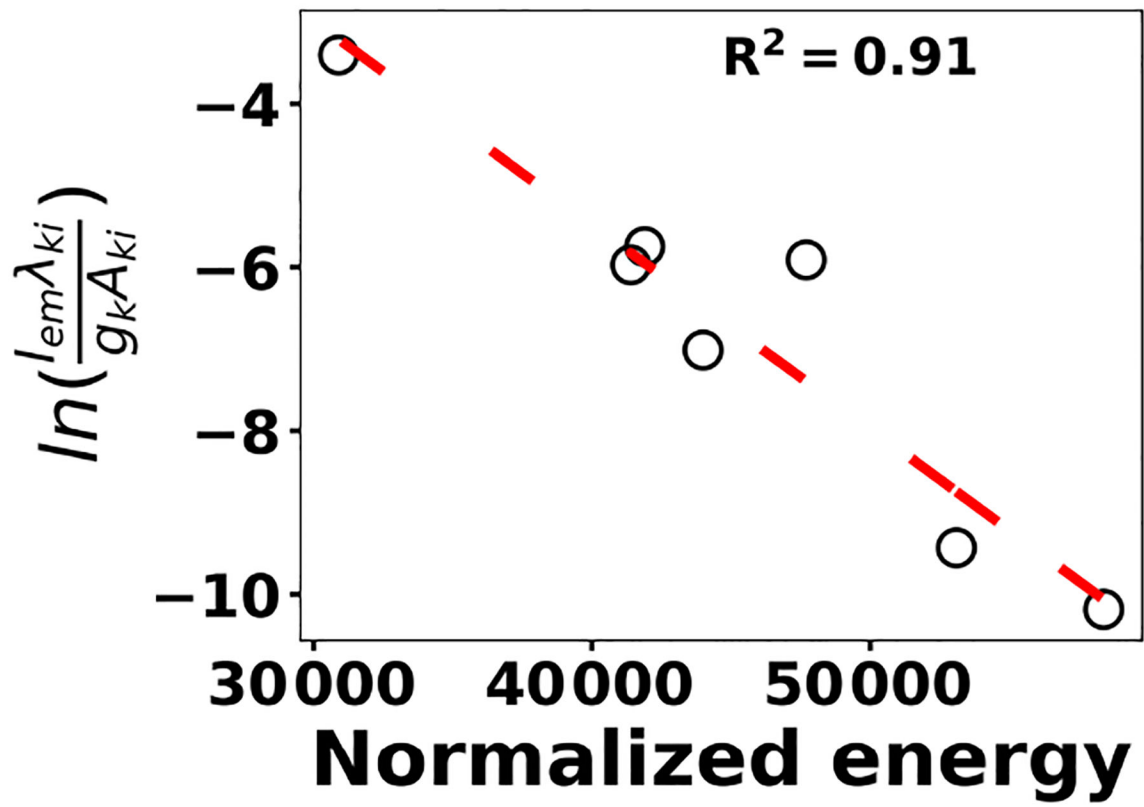


Figure 7. Boltzmann plot for various Tungsten lines in order to estimate plasma temperature.

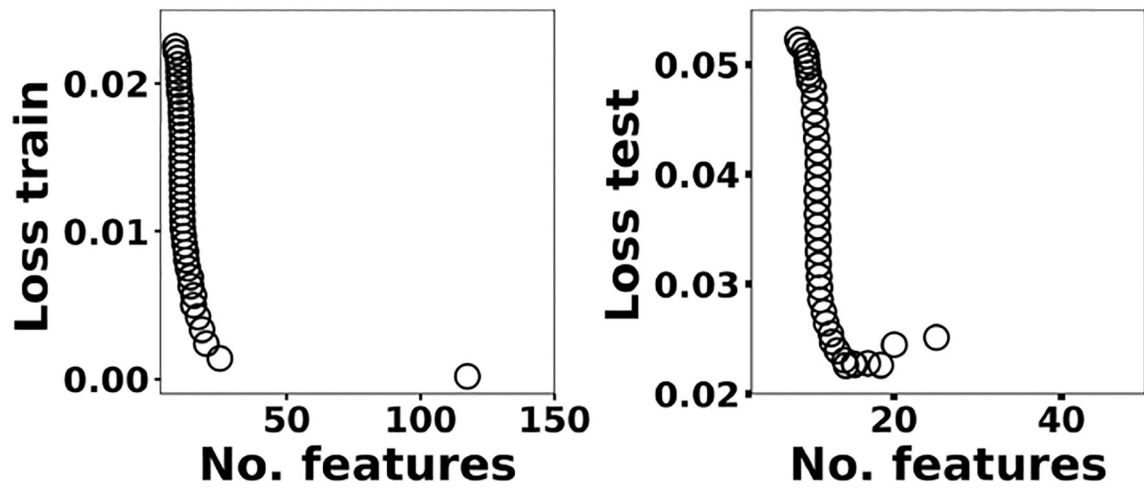


Figure 8.
The train and test losses for Ni as a function of the number of features.

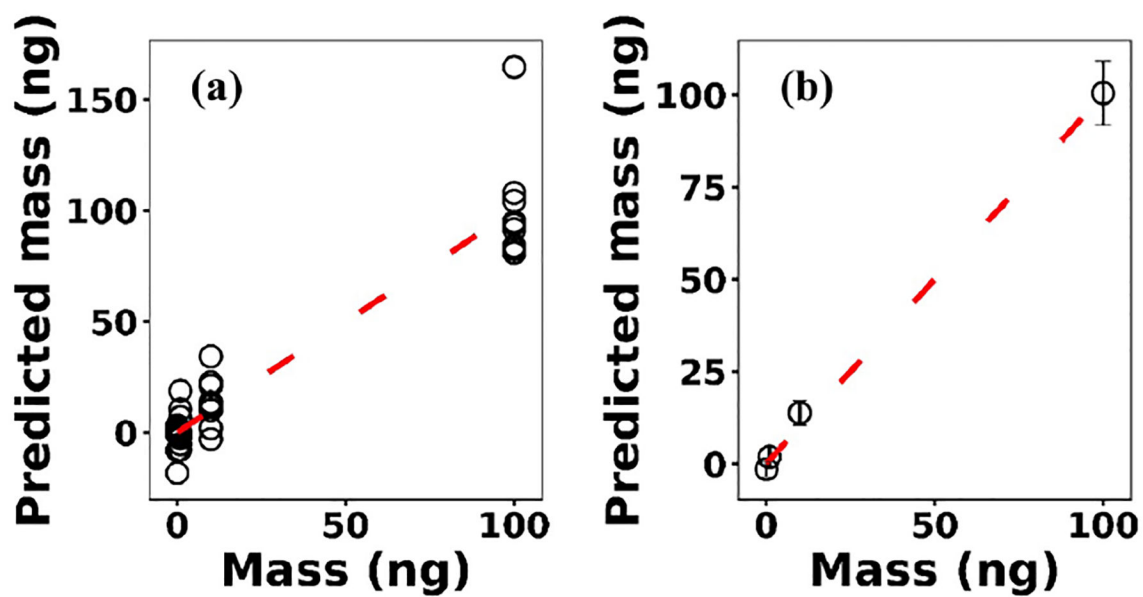


Figure 9.
(a) LASSO predictions based on leave-one-out cross validation for Ni; (b) the averaged predictions for each concentration.

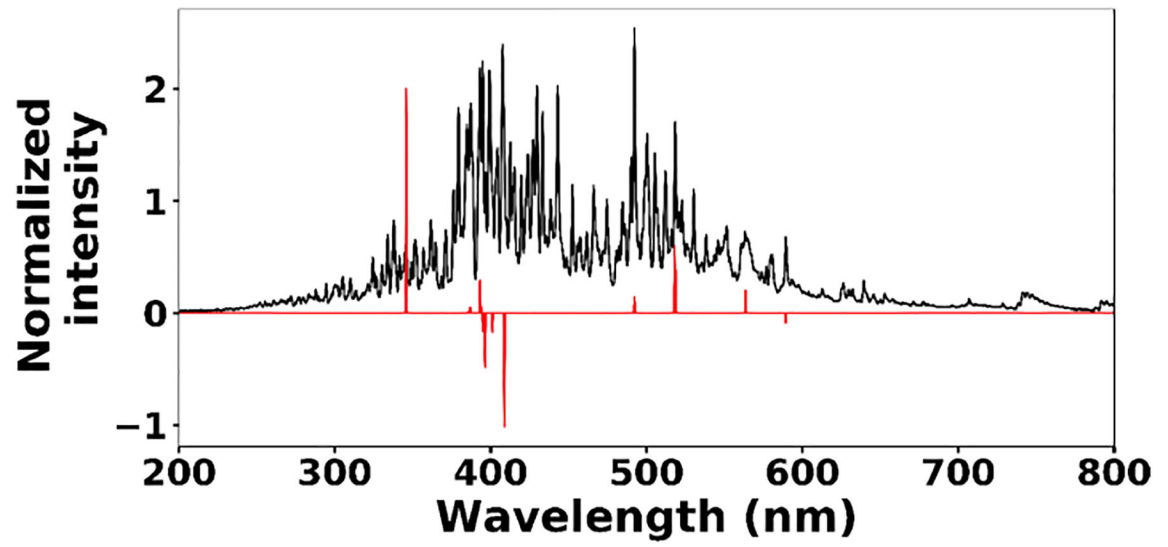


Figure 10.
Ni 10 ng spectrum (black line) and selected features by LASSO (red line).

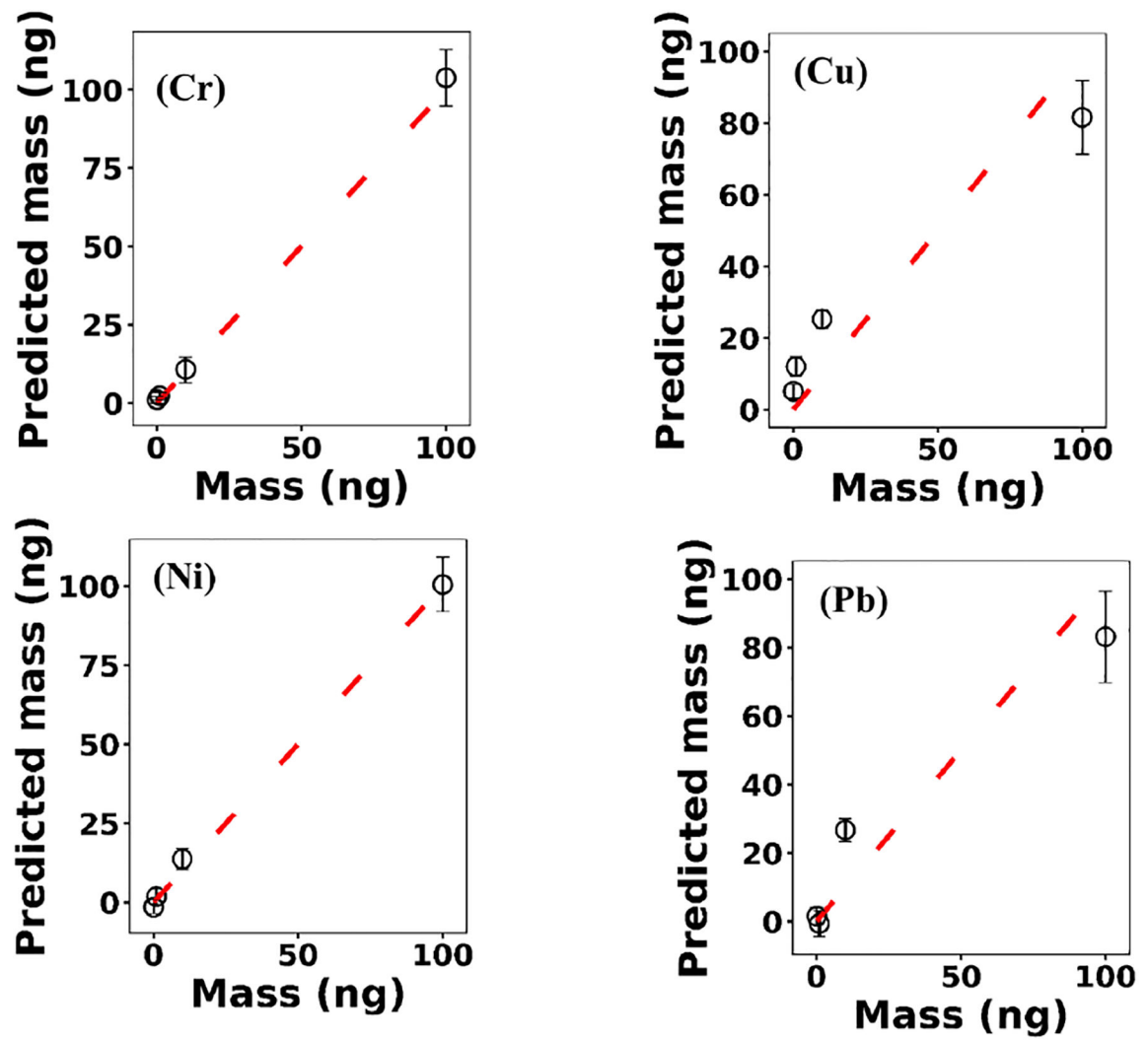


Figure 11.
The optimized LASSO models predictions for Cr, Cu, Ni, and Pb.

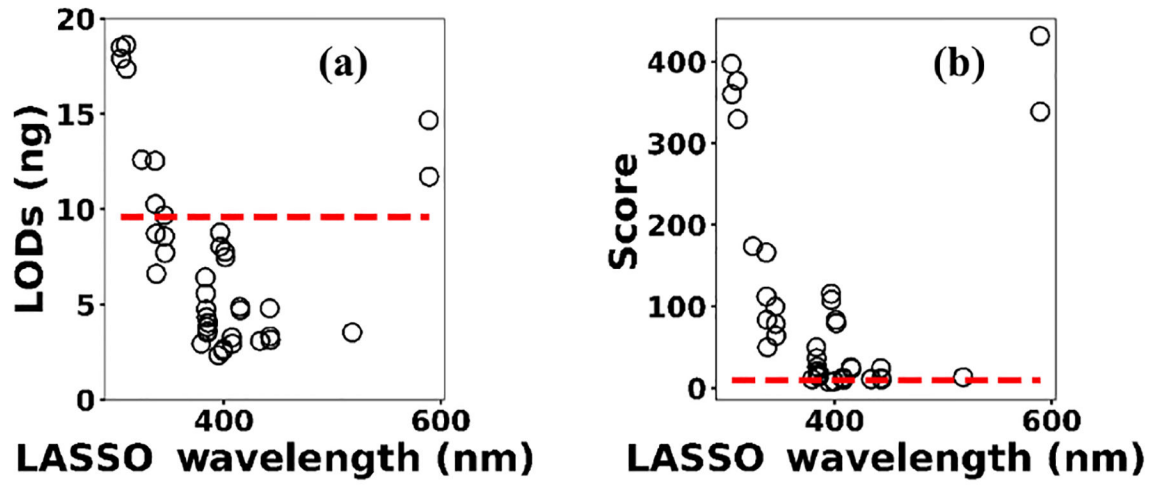


Figure 12.

(a) The univariate LODs based on LASSO selected features and (b) LASSO and univariate models scores.

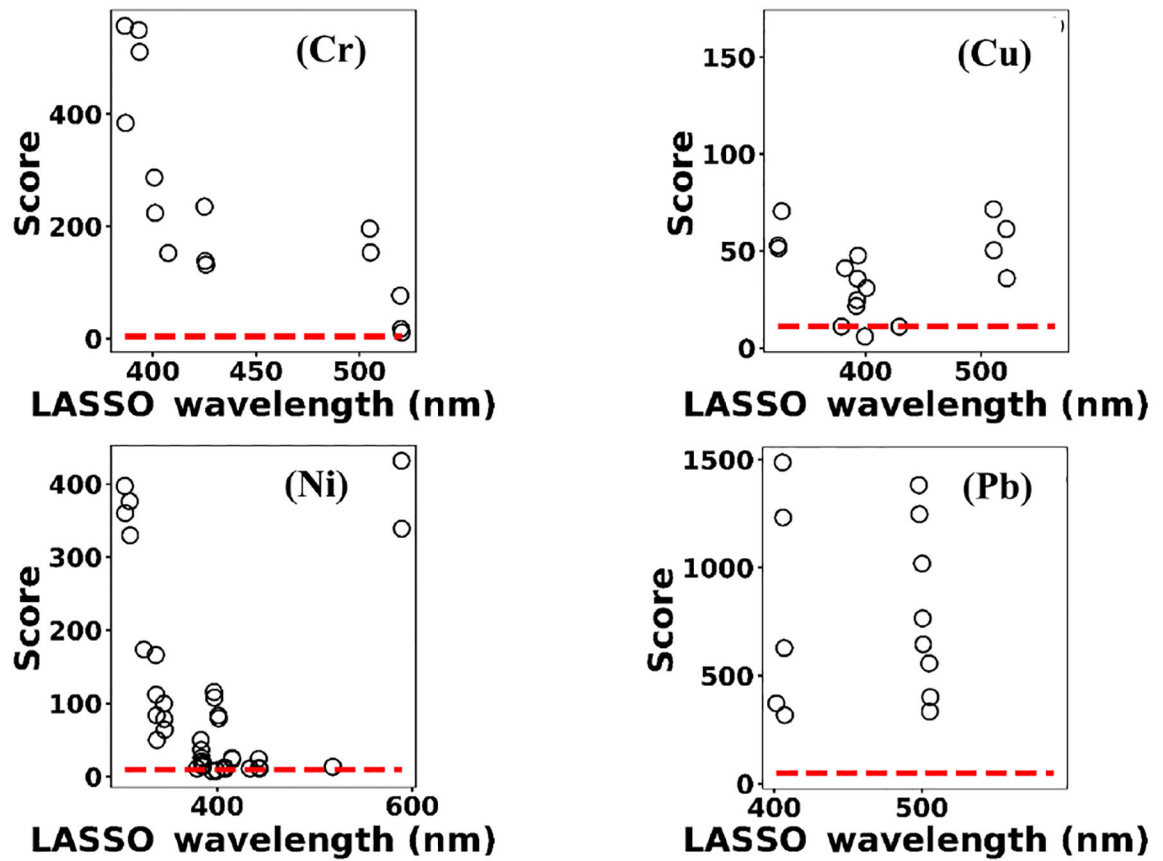


Figure 13. Model scores defined by Eq. (3) for Cr, Cu, Ni, and Pb. Circles indicate univariate models scores and dashed lines correspond to LASSO scores.

Table 1.

List of hazardous metals targeted by the US EPA.

US EPA metal HAPS
Antimony
Arsenic
Beryllium
Cadmium
Chromium
Cobalt
Lead
Manganese
Mercury
Nickel
Selenium

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

List of other toxic metals.

Toxic metals
Copper
Iron
Zinc

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Detection limits for various toxic metals based on the LASSO and univariate models.

Toxic metal	LASSO	R^2	MAE _{LASSO}	Univariate	R^2	MAE _{Univariate}	Regularization cons.
Cr	3.55	0.99	6.71	3.28	0.98	3.83	0.0008
Cu	12.09	0.92	49.67	0.68	0.11	143.27	0.0006
Ni	9.60	0.98	6.67	2.32	0.88	68.63	0.0009
Pb	54.40	0.90	36.67	8.37	0.45	124.42	0.0018