JAAS

PAPER

View Article Online
View Journal | View Issue

Cite this: *J. Anal. At. Spectrom.*, 2018, **33**, 404

Calibration approaches for the measurement of aerosol multielemental concentration using spark emission spectroscopy†

Lina Zheng,[ab] Pramod Kulkarni[*a] and Dionysios D. Dionysiou[b]

A multivariate calibration approach, using partial least squares regression, has been developed for the measurement of aerosol elemental concentration. A training set consisting of 25 orthogonal aerosol samples with 9 factors (elements: Cr, Mn, Fe, Ni, Cu, Zn, Cd, Pb, and Ti) and 5 levels (elemental concentrations) was designed. Spectral information was obtained for each aerosol sample using aerosol spark emission spectroscopy (ASES) at a time resolution of 1 minute. Simultaneously, filter samples were collected for the determination of elemental concentration using an inductively coupled plasma mass spectrometry (ICP-MS) analysis. Two regression models, PLS1 and PLS2, were developed to predict mass concentration from spectral measurements. The prediction ability of the models improved substantially when only signature wavelengths were included instead of the entire spectrum. The PLS1 model with 45 selected spectral variables (PLS1-45 model) presented the lowest relative root mean square error of cross validation (RMSECV; 16–35%). The detection limits using the PLS1-45 model for the nine elements were in the range of 0.16–0.50 μg m^{-3}. The performance of both multivariate and univariate regression models was tested for an unknown sample of welding fume aerosol. The multivariate model did not provide significantly better prediction compared to the univariate model. In spite of the difference in the matrices of the calibration aerosol and the unknown test aerosol, the results from the PLS model show good agreement with those from filter measurements. The relative root mean square error of prediction (RMSEP) obtained from the PLS1-45 model was 13% for Cr, 23% for Fe, 22% for Mn and 12% for Ni. The study shows that in spite of lower spectral resolution and lack of sample preparation, reliable and robust measurements can be obtained using the proposed calibration method based on PLS regression.

Received 21st July 2017
Accepted 23rd January 2018

DOI: 10.1039/c7ja00252a

rsc.li/jaas

1 Introduction

Exposure to airborne particles containing toxic metals in industrial atmospheres has adverse health effects on workers, thereby underscoring the need for reliable measurement methods for exposure characterization. The most widely used methods for determining concentrations of particulate toxic metals involve filter collection over several hours, followed by off-line analysis.[1,2] These methods are time-consuming and cannot capture short-term exposures. Workers cannot obtain instant feedback on their exposure to inhalable hazards. Developing real-time methods for measuring the aerosol chemical composition is of great significance to instantaneously assess exposure risks.

Our group has investigated various field portable microplasma spectroscopy techniques for real-time measurement of elemental species in aerosol particles[3–6] using broadband liquid crystal display (LCD) spectrometers. These techniques include laser-induced breakdown spectroscopy (LIBS), spark emission spectroscopy (SES) and glow discharge optical emission spectroscopy (GD-OES), and allow simultaneous multielemental measurement of aerosols. While these methods offer several practical advantages with respect to the development of low-cost, hand-portable instruments for exposure monitoring, they also suffer from drawbacks related to relatively inferior analytical figures of merit (compared to laboratory instruments) owing to a lack of optimal sample preparation and reduced wavelength resolution of LCD spectrometers. The objective of this study was to investigate the usefulness of multivariate calibration approaches to improve the calibration of microplasma methods for multielemental aerosol analysis.

For quantitative analysis of aerosols using microplasma spectroscopy, a univariate calibration approach has been widely used to construct calibration curves for relating the signal intensity and elemental concentration or mass.[3,4,7–10] However, in most spectroscopic methods, quantitative spectral analysis

[a]*Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Cincinnati, OH 45226, USA. E-mail: PSKulkarni@cdc.gov; Fax: +1-513-841-4545; Tel: +1-513-841-4300*

[b]*Environmental Engineering and Science Program, Department of Biomedical, Chemical and Environmental Engineering, University of Cincinnati, OH, 45221, USA*

† Electronic supplementary information (ESI) available: Univariate calibration curves and PLS regression coefficient plots. See DOI: 10.1039/c7ja00252a

remains challenging due to sample matrix effects (physical and chemical matrix effects) and spectral interference.[11] The emission signal intensity from one element is affected, among other factors, by the overall chemical composition of the sample. These matrix effects can result in large uncertainties if univariate calibration methods are used, especially when unknown samples have different chemical matrices from the calibration standards.[12–14] To minimize the sample matrix effects, matrix-matched standards are often used for univariate calibration in spectroscopic methods.[15,16] However, preparation and certification of matrix-matched standards are tedious for a given analytical measurement.

The application of multivariate calibration in spectroscopic analysis has proved to be beneficial in accounting for matrix effects as well as eliminating spectral interference.[13,17] Multivariate calibration has been widely used for the analysis of natural samples (such as rubber antioxidants, diesel fuel, seeds, and wine fermentations) by near-infrared (NIR) spectroscopy, which usually presents relatively weak and highly overlapping spectral bands.[18–22] Multivariate calibration has also been employed in atomic emission spectroscopy for elemental analysis of mixtures, such as steels, alloys, coal, and soil.[23–27] Unburned carbon in fly ash was analysed using LIBS and showed that multivariate calibration had better performance than univariate calibration as matrix effects caused by other components in fly ash could be taken into account.[13] Comparison of single and multivariate calibration for the determination of Si, Mn, Cr and Ni in highly alloyed stainless steels using LIBS has shown that multivariate analysis of spectral data was more effective and accurate when analytical lines overlapped.[17]

In this paper, we present a multivariate calibration approach for simultaneous measurement of multiple elements in aerosols using SES. Partial least squares (PLS) regression was used with a calibration training set to build PLS models for predicting the elemental concentration of aerosols. The prediction capability of the PLS models (PLS1 and PLS2) with different configurations was compared for an unknown test sample of welding fume aerosol.

2 Methods

2.1 Instrumentation

Spark emission spectroscopy was used for quantitative measurement of the elemental concentration of aerosols. Details of this method are described elsewhere.[4,5,28] Briefly, this method involves the collection of particles onto a small electrode tip (500 μm in diameter) with a corona aerosol microconcentrator (CAM),[28] followed by ablation of particles by a spark produced using a high voltage (HV) pulse generator (200 mJ per pulse; ARC-2, Cascodium Inc., Andover, MA). The optical emission from excited atomic and ionic species in the spark-induced plasma was collected using a broadband spectrometer (200–900 nm wavelength range and 0.1 nm resolution; LIBS 2500 Plus; Ocean Optics Inc.; Dunedin, FL) for spectrochemical analysis. A delay time of 5 μs and a gate width of 1 ms were used. The spectral data were used to identify elements and determine their mass in the collected particle samples.

2.2 Calibration aerosols

Aerosol containing metals of interest were generated from solutions using a pneumatic atomizer (model 3076, TSI Inc., Shoreview, MN). The precursor solutions used in the atomizer were prepared using water-soluble nitrates of the analytes of interest. A multilevel multifactor training set, containing various metals at various concentration levels, was designed for multivariate calibration.[29] It consisted of 25 samples with 9 mutually orthogonal factors (*i.e.*, elements: Cr, Mn, Fe, Ni, Cu, Zn, Cd, Pb, and Ti) and 5 levels (elemental concentrations in solution), as shown in Table 1.

2.3 Experimental procedure

Fig. 1 shows the schematic diagram of the multivariate calibration experimental procedure for the measurement of aerosol elemental concentrations. The test aerosol was generated using a pneumatic atomizer, and then passed through a diffusion dryer. The dry aerosol was introduced into a CAM for near real-time analysis using SES. The particles were collected in the CAM for 1 minute at a flow rate of 2 L min^{-1}, followed by SES measurement following procedures described elsewhere.[4,5] The test aerosol sample was analysed by SES in five consecutive measurements (the relative standard deviation of signal intensity was less than 15%), and their average was used to obtain the spectrum for each sample. The same test aerosol was also simultaneously collected on a polyvinyl chloride (PVC) filter for subsequent off-line elemental analysis. The particles were

Table 1 Elemental composition and concentration of solutions used for generating the calibration aerosol

	Elemental concentration in the solution, mg L^{-1}								
Sample no.	Cr	Mn	Fe	Ni	Cu	Zn	Cd	Pb	Ti
1	20	20	20	20	20	20	20	20	20
2	20	5	10	5	40	40	20	10	40
3	5	10	5	40	40	20	10	40	10
4	10	5	40	40	20	10	40	10	30
5	5	40	40	20	10	40	10	30	30
6	40	40	20	10	40	10	30	30	20
7	40	20	10	40	10	30	30	20	40
8	20	10	40	10	30	30	20	40	30
9	10	40	10	30	30	20	40	30	40
10	40	10	30	30	20	40	30	40	5
11	10	30	30	20	40	30	40	5	5
12	30	30	20	40	30	40	5	5	20
13	30	20	40	30	40	5	5	20	30
14	20	40	30	40	5	5	20	30	5
15	40	30	40	5	5	20	30	5	30
16	30	40	5	5	20	30	5	30	10
17	40	5	5	20	30	5	30	10	10
18	5	5	20	30	5	30	10	10	20
19	5	20	30	5	30	10	10	20	5
20	20	30	5	30	10	10	20	5	10
21	30	5	30	10	10	20	5	10	5
22	5	30	10	10	20	5	10	5	40
23	30	10	10	20	5	10	5	40	40
24	10	10	20	5	10	5	40	40	20
25	10	20	5	10	5	40	40	20	10

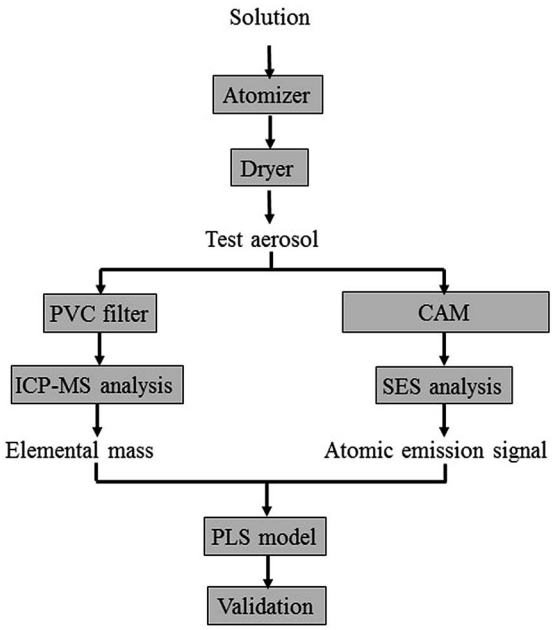

Fig. 1 A schematic diagram of the multivariate calibration experimental procedure for the measurement of aerosol elemental concentrations.

collected for 30 minutes at a flow rate of 2 L min^{-1}. Each filter sample was digested following NIOSH method 7303 (hot block/HCl/HNO_3 digestion) and then analysed using inductively coupled plasma mass spectrometry (ICP-MS). ICP-MS was used for filter sample analysis in this study because of its superior detection limits compared to ICP-AES for most elements. Very low detection limits were desirable in this study due to the short collection times (for filter samples) used in this study. Studies have shown that ICP-MS has far superior recoveries (~100%) and precision (<0.2) for the determination of metals in filter samples.[30] The elemental masses of the aerosol particles collected on the electrode in the CAM were deduced from the elemental air concentrations obtained from filter collection and ICP-MS analysis, and are shown in Table 2. A model was constructed using PLS regression as described below.

2.4 PLS regression model

A PLS model was constructed by using the measured spectra for all the test aerosols generated from the 'training set' calibration solutions. The model involves a linear regression between predictor/explanatory variables λ_i and response/dependent variables m_j. For each measurement, the predicted mass for element j, m_j, is given by

$$m_j = \sum_{i=0}^{p} \beta_{ji}\lambda_i + \varepsilon_j \quad (1)$$

where λ_i is the signal intensity at wavelength i of the emission spectra, β_{ji} is the linear regression coefficient for element j at each wavelength of the emission spectra and ε_j is a noise term for element j. p is the number of selected discrete wavelengths in each of the emission spectra.

Table 2 Elemental mass of the aerosol samples collected by the CAM determined by ICP-MS

	Elemental mass, ng								
Sample no.	Cr	Mn	Fe	Ni	Cu	Zn	Cd	Pb	Ti
1	21.0	21.5	19.5	22.5	25.5	19.0	20.0	20.5	17.5
2	18.5	5.0	8.8	4.3	37.0	34.0	18.0	8.5	32.0
3	6.0	9.5	4.4	40.0	36.0	15.5	8.5	35.5	8.0
4	10.5	5.0	39.5	43.0	21.0	8.5	39.0	12.0	27.0
5	6.0	41.0	45.5	22.5	10.5	36.0	9.5	30.0	28.0
6	37.5	38.5	18.6	9.5	37.0	8.5	28.5	29.5	18.0
7	38.5	20.5	9.5	41.5	10.0	23.5	29.0	20.0	34.5
8	19.5	10.5	32.5	10.0	28.0	23.5	19.5	38.5	26.5
9	10.0	38.0	9.2	30.0	28.0	16.0	38.0	29.0	32.0
10	38.5	10.5	24.5	31.5	20.0	35.0	29.5	39.0	2.6
11	11.5	30.0	29.0	22.0	38.5	25.0	39.0	6.0	5.0
12	31.5	32.0	19.0	44.0	30.0	36.0	6.0	5.5	18.5
13	30.5	21.0	36.0	33.0	40.0	5.0	5.0	19.5	26.0
14	18.0	36.5	26.5	39.5	5.0	4.5	18.5	27.5	3.9
15	39.5	30.0	41.5	5.5	5.0	16.5	29.5	5.0	28.0
16	27.5	37.5	4.3	4.0	17.5	23.5	4.7	26.5	8.5
17	34.0	4.6	4.1	18.5	25.0	3.1	25.5	9.0	7.5
18	5.5	4.7	22.0	28.5	4.7	23.0	8.5	8.0	11.0
19	5.5	20.0	23.5	4.8	27.5	8.5	9.0	19.0	4.3
20	17.5	27.0	4.4	29.0	8.5	8.0	18.0	4.8	7.5
21	27.5	4.9	26.0	8.5	8.5	16.0	4.5	8.5	4.5
22	6.0	29.5	9.2	9.0	19.0	3.9	9.0	5.0	36.5
23	29.0	10.0	9.6	20.0	6.0	8.0	4.8	36.0	35.0
24	9.5	8.5	19.0	3.7	7.5	2.8	32.0	34.0	16.0
25	10.0	20.0	4.9	9.5	5.5	35.5	40.5	21.5	9.0

The general matrix equation for the entire data takes the following form:

$$\begin{aligned} \boldsymbol{\lambda}_{np} &= \mathbf{TP}^{\mathrm{T}} + \boldsymbol{\varepsilon}_1 \\ \mathbf{m}_{nq} &= \mathbf{UQ}^{\mathrm{T}} + \boldsymbol{\varepsilon}_2 \end{aligned} \quad (2)$$

where $\boldsymbol{\lambda}_{np}$ is an $n \times p$ matrix of predictor variables, $\mathbf{m}_{nq}$ is an $n \times q$ matrix of response variables; $\mathbf{T}$ and $\mathbf{U}$ matrices represent the projections of $\boldsymbol{\lambda}_{np}$ and $\mathbf{m}_{nq}$, respectively; $\mathbf{P}$ and $\mathbf{Q}$ represent the matrices of loadings of the original variables; and $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are the residual matrices related to the noise. In this study, $n = 25$, corresponding to the number of total samples in the training set. The number of explanatory variables p is the same as the number of wavelengths included in the emission spectra. Also the number of response variables q is the same as the number of elements in each sample in the training set, and $q = 9$.

Two types of models were constructed for each analyte: model PLS1 and model PLS2. In model PLS1, only one element was modelled each time (*i.e.* nine separate PLS1 models were built for each element), while all nine elements were simultaneously used in a single PLS2 model.[31] Different numbers of explanatory variables (*i.e.* signal intensity) were selected as model inputs to build various models. The number of latent variables for each model was chosen to correspond to the lowest root mean square error of cross validation (RMSECV),[32] which typically ranged from 3 to 6 ng. Cross validation, which involves verifying the prediction capability of the PLS model, was performed using the 'leave-one-out' method. In this method, only one sample at a time is left out of the calibration and the

remaining samples are used for constructing a model to predict the sample that has been left out.[32] This procedure is repeated until all samples have been left out once. The performance of the regression models was evaluated by examining their relative RMSECV and correlation coefficients R^2 of cross validation. The PLS analyses were carried out using the software The Unscrambler 10.0 X (CAMO Software Inc., Woodbridge, NJ).

3 Results and discussion

3.1 Wavelength selection

Several studies involving chemometric analysis of spectroscopic data have employed select wavelengths as the explanatory variables, and have shown that the prediction ability of multivariate regression models could be improved when using the wavelengths that capture or explain maximum variance in the analyte concentration.[33,34] In this study, PLS regression coefficients were used for the selection of the most relevant wavelengths for each Y response variable (*i.e.*, elemental mass).[33–35] The wavelengths that have larger regression coefficients represent emission signals that are better correlated to the mass of the analyte.

Fig. 2 shows the PLS1 regression coefficients for Cr in the wavelength range of 300–540 nm. The highest regression coefficients were observed at a wavelength of around 520 nm. The magnified inset (Fig. 2(c)) presents three overlapping peaks at around 520 nm. These three peaks are consistent with the Cr atomic emission lines Cr I 520.45 nm, Cr I 520.60 nm, and Cr I 520.84 nm, according to the National Institute of Standards and Technology (NIST) atomic spectral database. In addition to the triplet peak, there are three consecutive peaks observed at around 360 nm (Fig. 2(a)), 425 nm (Fig. 2(b)) and 530 nm (Fig. 2(d)), respectively. They correspond to Cr I emission lines at 357.87 nm, 359.35 nm, 360.53 nm, 425.43 nm, 427.48 nm, 428.97 nm, 529.82 nm, 532.83 nm and 534.58 nm according to the NIST atomic spectral database. Similar relevant wavelengths for all the other elements of interest in this study were chosen by analysing their PLS regression coefficients. The PLS1 regression coefficients for all the remaining elements are shown in Fig. S1 in the ESI.† Five explanatory variables for each emission peak (a central wavelength pixel, with two additional pixels on each side) were used. For each element, wavelengths with regression coefficients greater than three times the standard deviation (3σ) around the mean regression coefficient (over all the wavelengths) were selected for subsequent verification. Table 3 shows the emission lines we have verified for each element based on the NIST atomic spectra database (ASD). These lines are not identified as self-absorbing or self-reversing lines in the ASD. The emission lines highlighted in bold were chosen to construct PLS1-45 and univariate regression models for each element; these lines exhibit relatively higher regression coefficients and minimal spectral interference.

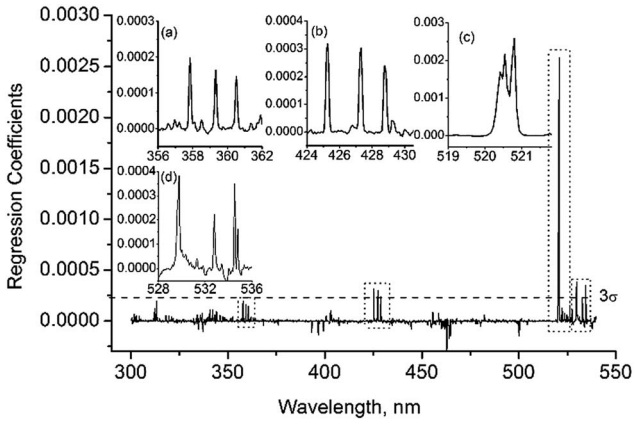

Fig. 2 PLS regression coefficients for chromium by analysis of the full range spectra (300–540 nm).

3.2 Multivariate calibration using PLS

Both PLS1 and PLS2 models were built by constructing the relationship between elemental mass and spectra. For constructing the PLS1 model, only one analyte was chosen as the response variable, $\mathbf{m}_{nj}$ ($j = 1, 2, \ldots,$ or 9). Here, nine PLS1 models were constructed separately for each analyte (*i.e.* response variable). For the PLS2 model, all nine response variables $\mathbf{m}_{nq}$ were chosen as the response variables simultaneously. Three $\boldsymbol{\lambda}_{np}$ matrices with different numbers of variables were selected as explanatory variables to build various PLS models. These three $\boldsymbol{\lambda}_{np}$ matrices include (i) explanatory variables containing the whole spectra ranging from 300 to 540 nm (4312 variables), (ii) significant explanatory variables containing multiple emission peaks related to the analytes (all the emission peaks shown in Table 3, *i.e.*, 290 variables), and (iii) a subset of significant explanatory variables containing only one emission peak for each element (*i.e.*, the ones with relatively higher regression coefficients and minimal spectral interference with other elements, highlighted in Table 3; 45 variables in total for all nine elements). To simplify the model description, these models were denoted as PLS1-4312, PLS1-290, PLS1-45, PLS2-4312, PLS2-290, and PLS2-45, respectively.

Table 3 Selected elemental emission lines for each element investigated in this study[a]

Elements	Emission lines/nm
Cr	425.43, 427.48, 428.97, 520.45, 520.60, 520.84, 529.82, **534.58**
Mn	344.20, 346.03, 347.40, 348.29, **403.08**, 403.31, 403.45, 482.35
Fe	374.56, 404.58, 406.36, 438.35, **526.95**
Ni	300.25, 341.48, 344.63, 345.29, 345.85, 346.17, 349.29, 351.03, 351.51, **352.45**, 356.64, 361.94
Cu	510.55, **521.82**
Zn	468.01, 472.22, **481.05**, 491.16, 492.40
Cd	361.05, 467.82, 479.99, 508.58, 533.75, **537.81**
Pb	**405.78**, 424.49, 438.65
Ti	334.94, 338.03, 338.38, 375.93, 453.32, 453.48, **498.17**, 499.11, 499.95

[a] The emission lines highlighted in bold are chosen to construct the PLS1-45 regression model and univariate regression model for each element.

Table 4 shows the relative RMSECV and R^2 obtained from cross validation for different models. Comparison of the PLS models constructed with different numbers of variables shows that the models with selected variables (PLS1-290, PLS1-45, PLS2-290, and PLS2-45) were significantly superior to the models with the whole spectra (PLS1-4312 and PLS2-4312). For PLS1 models, as the variable number is reduced from 4312 to 45, the average relative RMSECV for the nine elements is reduced from 36% to 25% and the average R^2 is improved from 0.65 to 0.84. This result demonstrates that variable selection plays an important role in multivariate calibration for elemental mass using spark emission spectroscopy. Although the PLS model has the ability to deal with numerous variables, the error of prediction could deteriorate if the entire raw spectral data are used in the calibration, because the majority of data consist of noise. Therefore, the use of selected explanatory variables that are most relevant to the underlying phenomenon as input variables is essential to construct the optimal PLS model. Comparison of PLS1 models using the selected wavelengths (PLS1-290 and PLS1-45) shows that their prediction abilities are quite similar. This is because the signal intensities of multiple emission peaks from the same element are correlated. Selection of one emission peak for each element as input variables of the PLS model is sufficient for constructing accurate multivariate models. Of all nine elements, Fe and Pb show low R^2 values, likely due to the high detection limits of Fe and Pb in our system (the detection limits are shown in Table 5). In this study, the elemental concentrations of Fe and Pb in the calibration aerosol were close to their limit of quantification, leading to increased measurement uncertainty and poor R^2.

Table 4 also shows that the relative RMSECVs of PLS1 models are 2–4% smaller than those of PLS2 models. This suggests that the PLS1 models have lower error than the PLS2 models. However, the need to construct a separate model for each element is a drawback of the PLS1 approach. On the other hand, one PLS2 model can be applied simultaneously to all the elements and is easier to implement.

Using the 25 multielement samples in the training set, a univariate regression model (U-ME) was also constructed for each element by plotting the signal intensity of an emission line as a function of elemental mass. For each element, the emission line with the highest regression coefficient was used. The signal intensity of an emission line was calculated as either the peak height or the peak area, both after baseline correction. The U-ME models were described using linear regression curves. These univariate calibration curves (based on peak height) are shown in Fig. S2 in the ESI.† The R^2 and RMSE of the univariate regression are also shown in Table 4. The R^2 values in parentheses are obtained from univariate regression using the peak area for each element. Both the peak height- and peak area-based calibration curves yielded similar sensitivities. For most elements, the U-ME model gives similar R^2 and RMSE values to those of the PLS models with select variables (*i.e.* PLS1-290, PLS1-45, PLS2-290, and PLS2-45). There is no consistent evidence in the literature about the effectiveness of different calibration methods (*i.e.* multivariate and univariate

Table 4 Comparison of relative RMSECV and R^2 from PLS1 and PLS2 models and those from U-ME models. The R^2 values for U-ME are for calibration curves based on peak height; those based on peak area are shown in parentheses

	R^2						
Element	PLS2-4312	PLS2-290	PLS2-45	PLS1-4312	PLS1-290	PLS1-45	U-ME
Cr	0.73	0.73	0.81	0.78	0.79	0.82	0.80(0.77)
Mn	0.66	0.71	0.76	0.73	0.78	0.82	0.75(0.77)
Fe	0.20	0.62	0.7	0.26	0.68	0.71	0.59(0.58)
Ni	0.58	0.74	0.75	0.66	0.84	0.83	0.71(0.73)
Cu	0.69	0.88	0.90	0.80	0.90	0.94	0.89(0.91)
Zn	0.60	0.83	0.80	0.68	0.79	0.84	0.83(0.83)
Cd	0.87	0.88	0.90	0.90	0.90	0.90	0.92(0.91)
Pb	0.15	0.47	0.76	0.20	0.64	0.81	0.84(0.83)
Ti	0.67	0.81	0.88	0.82	0.82	0.89	0.84(0.84)
Average	0.57	0.74	0.81	0.65	0.79	0.84	0.80(0.80)
	Relative RMSECV, %						
Element	PLS2-4312	PLS2-290	PLS2-45	PLS1-4312	PLS1-290	PLS1-45	U-ME
Cr	32	32	27	29	28	26	28
Mn	38	35	32	34	30	27	34
Fe	58	40	35	56	37	35	50
Ni	46	36	36	44	28	30	42
Cu	35	22	20	27	20	16	21
Zn	37	24	26	33	27	23	24
Cd	23	22	20	20	20	21	18
Pb	56	45	30	55	37	27	25
Ti	35	26	20	25	25	19	24
Average	40	31	27	36	28	25	30

Table 5 Limits of detection of aerosol spark emission spectroscopy

	Limits of detection			
	In terms of mass, ng		In terms of air concentration[a], μg m^{-3}	
Elements	Multivariate calibration (PLS1-45)	Univariate calibration (U-ME)	Multivariate calibration (PLS1-45)	Univariate calibration (U-ME)
Cr	1.5	1.8	0.28	0.36
Mn	1.2	1.8	0.24	0.36
Fe	2.5	3.9	0.50	0.78
Ni	2.0	3.5	0.40	0.70
Cu	1.1	1.3	0.22	0.26
Zn	1.8	2.1	0.36	0.42
Cd	0.8	1.3	0.16	0.26
Pb	2.4	4.0	0.48	0.80
Ti	1.1	1.1	0.22	0.22

[a] Assuming a flow rate of 2 L min^{-1} and a sample collection time of 5 min.

methods).[13,14,17] For example, some studies on sample analysis using LIBS suggest that multivariate calibration provides improved accuracy and precision compared with univariate calibration,[13,14] while another study on the determination of elements in stainless steels using LIBS showed that univariate calibration provided the best prediction if appropriate, non-overlapping emission lines could be found.[17] To a certain extent, the performance of different calibration approaches depends on both spectral overlap and design of the calibration experiments.[36] In this study, we designed the training set with orthogonal factors (using elements relevant to welding aerosol measurement) and selected the non-overlapping emission peaks for calibration. Both the univariate and multivariate calibration approaches showed similar prediction capabilities.

Fig. 3 shows a comparison of calibration values (open circles) with the corresponding cross-validation (solid triangles) for the PLS1-45 model. The mass predicted by the model is plotted on the *y*-axis and the corresponding measured mass on the *x*-axis. These models were constructed with 9, 5, 6, 5, 3, 4, 3, 6, and 5 latent variables for Cr, Mn, Fe, Ni, Cu, Zn, Cd, Pb, and Ti, respectively. The PLS regression results (open circles) show a strong correlation between the predicted and measured elemental mass, with R^2 values in the range of 0.81–0.96 for different elements. The cross validation results (solid triangles) also show a strong correlation with an average R^2 value of 0.84. The relative RMSECVs of PLS1-45 models for different elements are in the range of 16–35%. For Fe, a low R^2 (0.71) and high RMSECV (35%) were observed, probably due to the poor detection limits of Fe in our system (the detection limit is shown in Table 5). A large uncertainty might be produced when the elemental mass concentration of the calibration aerosol is near the limits of quantification. Matrix effects may also have been relatively more significant for Fe measurements.

3.3 Limits of detection

According to 3σ criteria defined by the International Union of Pure and Applied Chemistry (IUPAC),[37] the limit of detection (LOD) for univariate calibration is expressed as

$$m_{\mathrm{LOD}} = \frac{3\sigma}{S} \tag{3}$$

where σ is the standard deviation of the blank at the selected emission line and S is the sensitivity given by the slope of the calibration curve of univariate calibration. Blank measurements were taken and σ was obtained by averaging over 20 replicate blank measurements.

The LOD in multivariate calibration was determined analogously to the univariate calibration:[38,39]

$$m_{\mathrm{LOD}} = \frac{3.3\sigma}{S} = 3.3\sigma\|b\| = 3.3\sigma\sqrt{{b_1}^2 + \ldots + {b_n}^2} \tag{4}$$

where σ is an estimate of the noise level in the data, S is the sensitivity of the PLS model, which gives the fraction of analytical signal due to the increase of the concentration of a particular analyte at unit concentration, $\|b\|$ is the Euclidian norm of the regression coefficient vector, and b_n is the n^{th} regression coefficient. σ was obtained by measuring the variation of the noise in the selected regions for each analyte. S was estimated as the inverse of $\|b\|$. Table 5 shows the LOD obtained from both univariate calibration and multivariate calibration (PLS1-45 model) for elemental measurements using SES. The mass LOD for multivariate calibration is in the range of 0.8–2.5 ng depending on the element, with the lowest LOD for Cd and the highest LOD for Fe. The LOD for multivariate calibration is in the range of 0.16–0.50 μg m^{-3} in terms of air concentration, calculated by assuming a flow rate of 2 L min^{-1}, a collection time of 5 minutes, and a particle collection efficiency of 0.5 for our system. Comparison of the LODs from univariate and multivariate calibration shows that both approaches provide similar LODs for most elements (except Fe, Ni, and Pb). This result is consistent with that found by Braga *et al.* in their study on the comparison of univariate and multivariate calibration for the determination of micronutrients in plant materials using LIBS.[39] It is worth noting that the LOD for Fe, Ni, and Pb from univariate calibration is nearly twice as high as that from multivariate calibration. This difference could possibly be due to higher measurement uncertainties for these elements arising from matrix effects.

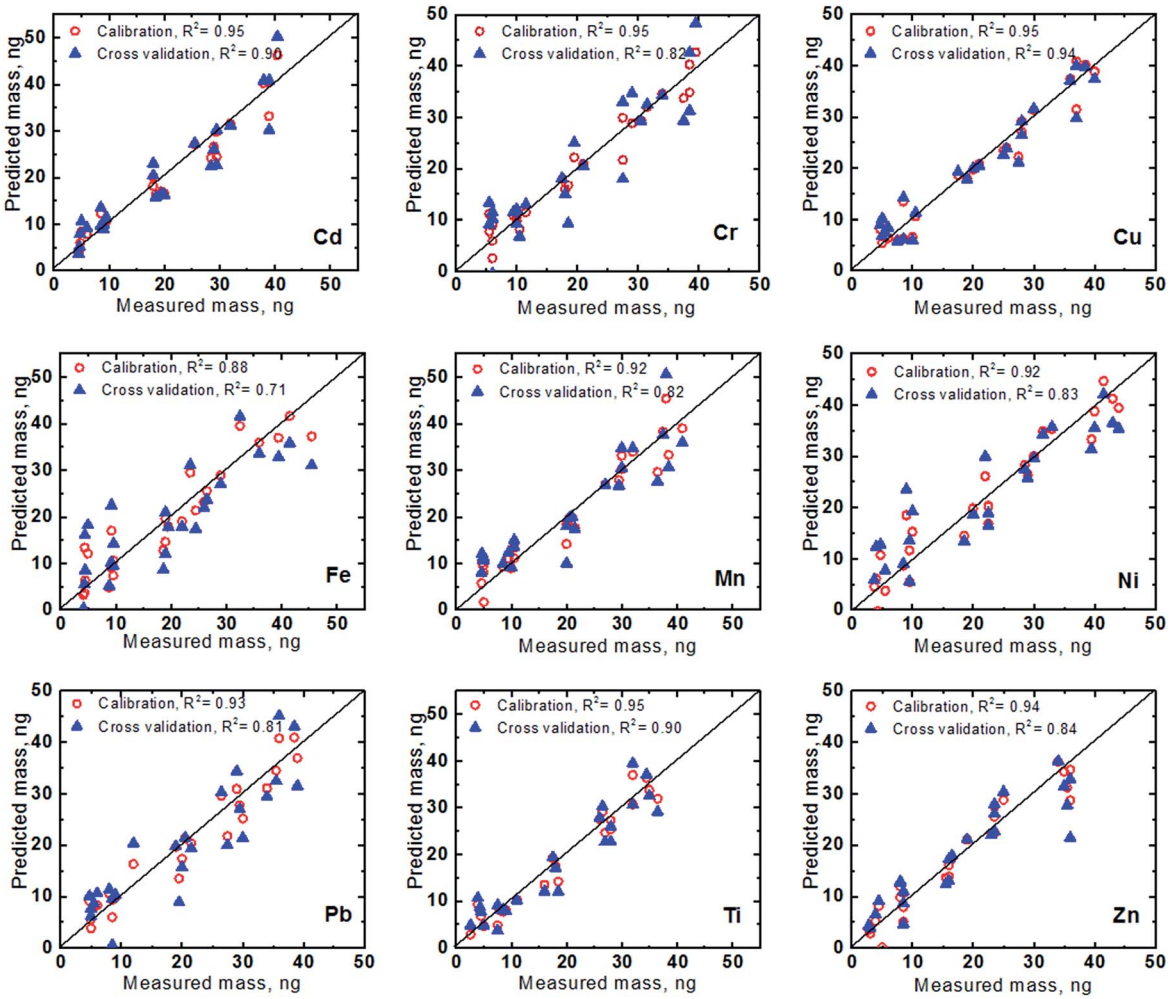

Fig. 3 Predicted *vs.* measured elemental mass in the collected particle samples for PLS1-45 models.

3.4 Application to welding aerosol measurement

We applied the multivariate regression models (PLS1-45 and PLS1-290), developed using the training set, to the measurement of the unknown elemental concentration of a welding aerosol. The test welding aerosol was generated by aerosolizing a suspension of stainless steel welding fume reference material (HSL SSWF-01, Health & Safety Laboratory, Buxton, UK). It should be noted that the matrix composition of this reference material (*i.e.*, the number of elements and their relative concentration in the matrix) and its solubility are different from those of the calibration samples (Table 1). The compounds used for calibration were soluble, while the welding fume particles were insoluble. From the suspension containing the welding fume reference material, aerosols with five different air concentration levels were obtained through dilution of the aerosol with clean, particle-free air. Test aerosols were collected for 2 minutes using an ASES system at 2 L min^{-1}. The elemental masses in the collected samples were predicted using both multivariate and univariate regression models, and then converted to elemental air concentrations. Three repeat measurements were performed by ASES for each test aerosol sample. Simultaneously, filter samples were collected (for 30 min) in parallel to obtain the elemental concentrations of the test welding aerosol by laboratory ICP-MS analysis after sample digestion. The test aerosols were relatively stable during these 30 minutes.

Fig. 4 shows a comparison of elemental concentrations predicted by the regression models: (a) PLS1-45 model, (b) PLS1-290 model, and (c) univariate calibration from single element solutions (U-SE model). Model predictions are compared with the corresponding ICP-MS measurements shown on the *x*-axis. Fig. 4 shows that most (17 out of 20 for the PLS1-45 model, 16 out of 20 for the PLS1-290 model, and 15 out of 20 for the U-SE model) samples agreed with the ICP-MS measurements within a ±25% bias. The results from the multivariate models [Fig. 4(a) and (b)] show good agreement with those from filter samples using ICP-MS. However, it is worth noting that the univariate calibration curve exhibits non-linear behaviour at higher mass loadings. The effect was not pronounced at the mass loadings studied in this work; however, the extrapolation of the curve seems to suggest significantly nonlinear behaviour at mass loadings above 100 μg m^{-3} [Fig. 4(c)]. In spite of the difference in matrices, with respect to solubility and elemental

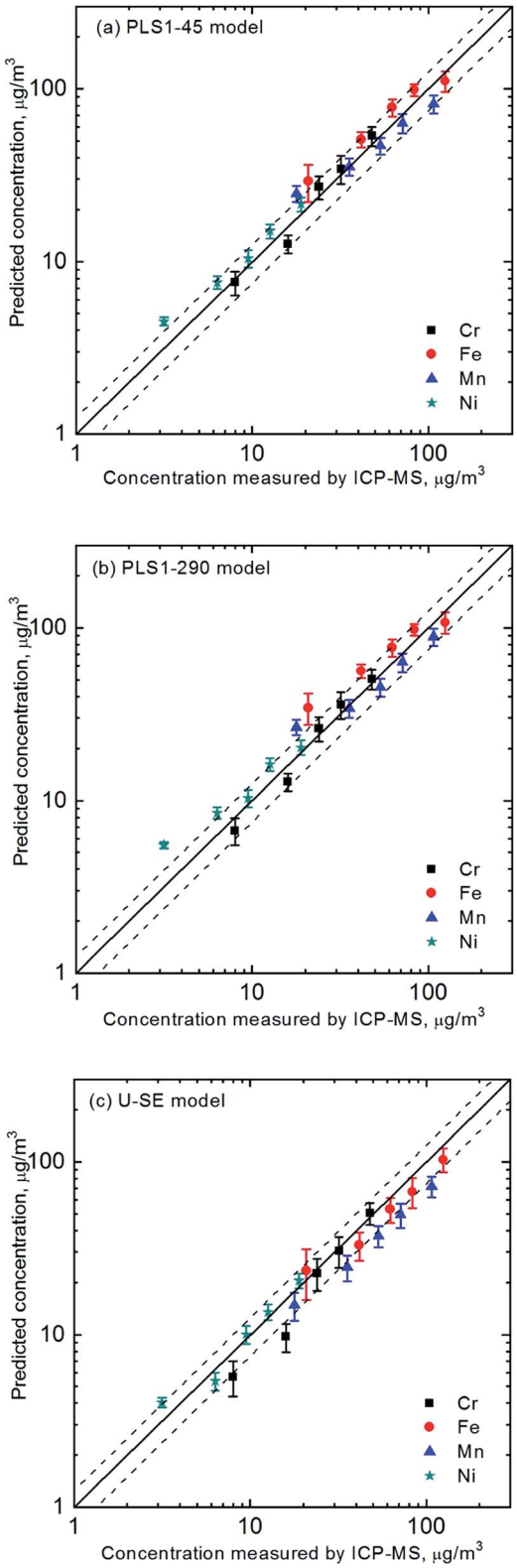

Fig. 4 Comparison of elemental concentrations predicted by the regression models: (a) PLS1-45 model, (b) PLS1-290 model and (c) U-SE model, and concentrations measured by ICP-MS analysis for welding aerosol measurement. The two dashed lines represent ±25% error around the 1 : 1 line.

composition, of the calibration aerosol and welding fume aerosol, the agreement between the ASES measurements and the ICP-MS measurements is good. This demonstrates that the prediction models constructed using particles aerosolized from soluble precursor species could be reliably used for the measurement of unknown insoluble (or soluble) particles using ASES. This significantly simplifies the calibration of field portable ASES instruments.

The relative root mean square error of prediction (RMSEP) was used to estimate the accuracy of our regression models. The relative RMSEP from the PLS1-45 model was 13% for Cr, 23% for Fe, 22% for Mn and 12% for Ni. The relative RMSEP from the PLS1-290 model was 12% for Cr, 25% for Fe, 21% for Mn and 13% for Ni. The relative RMSEP from the U-SE model was 14% for Cr, 27% for Fe, 29% for Mn and 10% for Ni. The multivariate calibration approach resulted in slightly better prediction accuracy than the single element univariate calibration approach when they were applied to the measurement of the welding fume aerosol. The welding aerosol represented a smaller subset of the calibration aerosol in terms of the number of elements—the calibration aerosol had nine elements, whereas the welding aerosol had only four elements. Moreover, as noted above, unlike the calibration aerosols, the welding aerosol particles were insoluble in water. The matrix effects in these samples may not be pronounced, which may explain the somewhat similar results obtained from both univariate and multivariate methods. However, the multivariate method may provide better precision in cases where significant matrix effects or nonlinear calibration curves are expected. Multivariate methods should also provide better precision and accuracy for cases involving significant spectral overlap or interference for the analytes of interest.

The error bars in Fig. 4 show the standard deviation of three replicate measurements, which ranged from 7–19%, demonstrating good precision of the ASES aerosol measurement system.

4 Conclusions

A multivariate calibration approach for the measurement of aerosol elemental concentrations by SES was developed. An orthogonal training set design (Table 1) in combination with suitable multivariate calibration allows one to make full use of multielemental emission spectra without the loss of spectral selectivity. It was shown that PLS regression coefficients can be judiciously used as a guide for selecting the most relevant portions of the emission spectra and to allow the construction of meaningful multivariate regression models.

Creating separate PLS models for individual elements (*i.e.* model PLS1) offers a slightly improved prediction ability than that provided by a single PLS model for multiple elements (model PLS2). Using a training set with orthogonal factors, it was found that multivariate models provided slightly better prediction compared to the univariate model. The simplicity of univariate calibration methods may be more attractive in applications where the unknown matrix has fewer elements and is relatively homogenous. However, multivariate calibration

may be preferable for aerosol particles containing large numbers of elements with highly varying elemental ratios. The multivariate model may also provide better precision when significant spectral interference or overlapping is expected.

Application of the PLS models to a welding aerosol also showed good agreement with ICP-MS measurements in spite of a significant difference in the calibration sample matrix (9 elements; soluble compounds) and unknown sample (4 elements; insoluble compounds). In spite of relatively lower spectral resolution and lack of sample preparation in our ASES method, the good agreement with the ICP-MS method suggests that using PLS regression with an orthogonal training set provides a relatively robust calibration method for field-portable microplasma methods for elemental analysis.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This research work is supported by intramural NORA and NTRC grants at NIOSH (CAN ZKPY, ZJLS, and ZLFX). The authors would like to thank Dr Kevin Ashley and Amy Feng for providing helpful feedback and review of this manuscript.

References

1 P. C. Raynor, D. Leith, K. Lee and R. Mukund, in *Aerosol Measurement: Principles, Techniques and Applications*, ed. P. Kulkarni, P. A. Baron and K. Willeke, John Wiley & Sons, Hoboken, New Jersey, 2011, ch. 7.
2 K. R. Spurny, *Analytical Chemistry of Aerosols: Science and Technology*, CRC Press, Boca Raton, Florida, 1999.
3 P. Diwakar, P. Kulkarni and M. E. Birch, *Aerosol Sci. Technol.*, 2012, **46**, 316–332.
4 P. K. Diwakar and P. Kulkarni, *J. Anal. At. Spectrom.*, 2012, **27**, 1101–1109.
5 L. Zheng, P. Kulkarni, M. E. Birch, G. Deye and D. D. Dionysiou, *Aerosol Sci. Technol.*, 2016, **50**, 1155–1166.
6 L. Zheng and P. Kulkarni, *Anal. Chem.*, 2017, **89**, 6551–6558.
7 M. Z. Martin, M. D. Cheng and R. C. Martin, *Aerosol Sci. Technol.*, 1999, **31**, 409–421.
8 A. J. R. Hunter, S. J. Davis, L. G. Piper, K. W. Holtzclaw and M. E. Fraser, *Appl. Spectrosc.*, 2000, **54**, 575–582.
9 M. Khalaji, B. Roshanzadeh, A. Mansoori, N. Taefi and S. H. Tavassoli, *Optic. Laser Eng.*, 2012, **50**, 110–113.
10 R. K. Marcus, M. A. Dempster, T. E. Gibeau and E. M. Reynolds, *Anal. Chem.*, 1999, **71**, 3061–3069.
11 P. Gemperline, *Practical Guide to Chemometrics*, CRC Press, Boca Raton, Florida, 2006.
12 S. Laville, M. Sabsabi and F. R. Doucet, *Spectrochim. Acta, Part B*, 2007, **62**, 1557–1566.
13 S. Yao, J. Lu, J. Zheng and M. Dong, *J. Anal. At. Spectrom.*, 2012, **27**, 473–478.
14 M. M. Tripathi, K. E. Eseller, F.-Y. Yueh and J. P. Singh, *Spectrochim. Acta, Part B*, 2009, **64**, 1212–1218.
15 D. A. Cremers and L. J. Radziemski, *Handbook of Laser-Induced Breakdown Spectroscopy*, John Wiley & Sons, Chichester, England, 2006, pp. 99–117.
16 M. da Silva Gomes, G. G. A. de Carvalho, D. Santos and F. J. Krug, *Spectrochim. Acta, Part B*, 2013, **86**, 137–141.
17 S. M. Zaytsev, A. M. Popov, E. V. Chernykh, R. D. Voronina, N. B. Zorov and T. A. Labutin, *J. Anal. At. Spectrom.*, 2014, **29**, 1417–1424.
18 C. E. Boschetti and A. C. Olivieri, *Journal of Near Infrared Spectroscopy*, 2001, **9**, 245–254.
19 M. C. Breitkreitz, I. M. Raimundo, J. J. R. Rohwedder, C. Pasquini, H. A. Dantas, G. E. Jose and M. C. U. Araujo, *Analyst*, 2003, **128**, 1204–1207.
20 T. A. Lestander and P. Geladi, *Can. J. For. Res.*, 2005, **35**, 1139–1148.
21 X. Shao, X. Bian, J. Liu, M. Zhang and W. Cai, *Anal. Methods*, 2010, **2**, 1662–1666.
22 M. Zeaiter, J. M. Roger and V. Bellon-Maurel, *Chemom. Intell. Lab. Syst.*, 2006, **80**, 227–235.
23 M. Z. Martin, N. Labbé, N. André, S. D. Wullschleger, R. D. Harris and M. H. Ebinger, *Soil Sci. Soc. Am. J.*, 2010, **74**, 87–93.
24 F. R. Doucet, T. F. Belliveau, J.-L. Fortier and J. Hubert, *Appl. Spectrosc.*, 2007, **61**, 327–332.
25 A. Golloch and K. Wilke, *J. Anal. At. Spectrom.*, 1997, **12**, 1225–1230.
26 F. B. Gonzaga and C. Pasquini, *Spectrochim. Acta, Part B*, 2012, **69**, 20–24.
27 J. Feng, Z. Wang, L. Li, Z. Li and W. Ni, *Appl. Spectrosc.*, 2013, **67**, 291–300.
28 L. Zheng, P. Kulkarni, K. Zavvos, H. Liang, M. E. Birch and D. D. Dionysiou, *J. Aerosol Sci.*, 2017, **104**, 66–78.
29 R. G. Brereton, *Analyst*, 1997, **122**, 1521–1529.
30 K. Ashley, S. A. Shulman, M. J. Brisson and A. M. Howe, *J. Environ. Monit.*, 2012, **14**, 360–367.
31 R. G. Brereton, *Analyst*, 2000, **125**, 2125–2154.
32 R. G. Brereton, *Applied Chemometrics for Scientists*, John Wiley & Sons, Chichester, England, 2007.
33 A. G. Frenich, D. Jouan-Rimbaud, D. L. Massart, S. Kuttatharmmakul, M. M. Galera and J. L. M. Vidal, *Analyst*, 1995, **120**, 2787–2792.
34 H. Xu, Z. Liu, W. Cai and X. Shao, *Chemom. Intell. Lab. Syst.*, 2009, **97**, 189–193.
35 S. D. Osborne, R. Künnemeyer and R. B. Jordan, *Analyst*, 1997, **122**, 1531–1537.
36 C. Demir and R. G. Brereton, *Analyst*, 1998, **123**, 181–189.
37 P. Boumans, *Anal. Chem.*, 1994, **66**, 459A–467A.
38 P. Valderrama, J. W. B. Braga and R. J. Poppi, *J. Agric. Food Chem.*, 2007, **55**, 8331–8338.
39 J. W. B. Braga, L. C. Trevizan, L. C. Nunes, I. A. Rufini, D. Santos and F. J. Krug, *Spectrochim. Acta, Part B*, 2010, **65**, 66–74.

