

Discrimination of biological and chemical threat simulants in residue mixtures on multiple substrates

Jennifer L. Gottfried

Received: 13 December 2010 / Revised: 27 January 2011 / Accepted: 28 January 2011 / Published online: 18 February 2011
© Springer-Verlag (outside the USA) 2011

Abstract The potential of laser-induced breakdown spectroscopy (LIBS) to discriminate biological and chemical threat simulant residues prepared on multiple substrates and in the presence of interferents has been explored. The simulant samples tested include *Bacillus atrophaeus* spores, *Escherichia coli*, MS-2 bacteriophage, α -hemolysin from *Staphylococcus aureus*, 2-chloroethyl ethyl sulfide, and dimethyl methylphosphonate. The residue samples were prepared on polycarbonate, stainless steel and aluminum foil substrates by Battelle Eastern Science and Technology Center. LIBS spectra were collected by Battelle on a portable LIBS instrument developed by A3 Technologies. This paper presents the chemometric analysis of the LIBS spectra using partial least-squares discriminant analysis (PLS-DA). The performance of PLS-DA models developed based on the full LIBS spectra, and selected emission intensities and ratios have been compared. The full-spectra models generally provided better classification results based on the inclusion of substrate emission features; however, the intensity/ratio models were able to correctly identify more types of simulant residues in the presence of interferents. The fusion of the two types of PLS-DA models resulted in a significant improvement in classification performance for models built using multiple substrates. In addition to identifying the major components of residue mixtures, minor components such as growth media and solvents can be identified with an appropriately designed PLS-DA model.

Keywords LIBS · PLS-DA · Biological warfare agents · Chemical warfare agents · Residue detection

Introduction

The development of reliable sensors to detect signs of terrorist activity is a high priority for the military and Homeland Security. In addition to the pressing need for explosives detection, the use of Sarin gas in the 1995 Tokyo subway attack and the distribution of anthrax spores through the US Postal Service in 2001 demonstrate the immediacy of the need for chemical and biological warfare agent sensors as well. Laser-induced breakdown spectroscopy (LIBS) has been investigated as a sensor for hazardous materials analysis. LIBS is a spectroscopic analysis technique that uses the light emitted from a laser-induced microplasma to determine the composition of the sample based on elemental and molecular emission intensities [1, 2]. LIBS provides real-time detection of solids, liquids, and vapors without the need for sample preparation and can be configured for standoff, laboratory close-contact, or portable detection. A number of groups have demonstrated the capability of LIBS for chemical [3–5], biological [6–18], and explosive detection [19]. With the exception of explosive residue detection, most of the studies on chemical and biological threat detection with LIBS have concentrated on the discrimination of bulk material (or on aerosolized particles, e.g., [20–23]).

Recent advances in chemometric analysis of LIBS data have enabled the discrimination of a variety of complex molecules and materials [7, 9, 16, 18, 19, 21, 24–27], despite the fact that LIBS is essentially an elemental analysis technique. Chemometric analysis

Published in the special issue *Laser-Induced Breakdown Spectroscopy* with Guest Editors Jagdish P. Singh, Jose Almirall, Mohamad Sabsabi, and Andrzej Miziolek.

J. L. Gottfried (✉)
US Army Research Laboratory, RDRL-WML-B,
Aberdeen Proving Ground,
Aberdeen, MD 21005, USA
e-mail: jennifer.gottfried@us.army.mil

techniques that have been applied to LIBS for discrimination include principal components analysis (e.g., [7, 25, 28, 29]), soft independent method of class analogy (e.g., [24, 25, 28]), discriminant function analysis (e.g., [12, 30]), and partial least-squares discriminant analysis (PLS-DA, e.g., [16, 19, 27]). Our group at the US Army Research Laboratory (ARL) has had the most success with PLS-DA, particularly for explosive residue detection applications.

PLS-DA is a multivariate inverse least-squares discrimination method used to classify samples [31, 32]. Predictor variables called latent variables (LV) are calculated based on linear combinations of the input variables in order to maximize the variance between sample classes. LIBS spectra typically have large intraclass variance due to shot-to-shot variability. Single-shot LIBS spectra of residues have even larger intraclass variability because the residue is often heterogeneously distributed on the substrate. A technique that maximizes the interclass variability, such as PLS-DA, is therefore important for good discrimination of residues with LIBS. The optimal number of LV for each PLS-DA model is determined by the point at which increasing the number of LV incorporates noise and other non-relevant information into the fit. Including too many LV could lead to overfitting in the model; overfitting can be avoided by testing the model with additional sample spectra not used to train the model [33]. The “variable importance in projection” (VIP) scores are calculated for each class and are essentially a weighting of the regression vectors showing the variables that are most important for separation between the classes in the model [34, 35]. A variable with a VIP score close to or greater than 1 can be considered important in a given model. Another output of the PLS-DA calculation is the Y predictor matrix used to estimate class affiliation. Unknown samples can be tested against the model, and the model will provide predictions about class designations. A threshold to determine whether a sample is in the class is established by the model using Bayesian statistics in order to minimize the number of false positives and false negatives. The model also calculates the percent predicted probability that a test sample belongs to each class in the model.

In this study, we applied PLS-DA techniques used to discriminate explosive residues on multiple substrates [36] to the problem of chemical and biological warfare agent simulant residue detection in the presence of interferents on multiple substrates. This is a particularly challenging problem for LIBS—in addition to discriminating complex chemical and biological materials using only atomic and molecular emission intensities, the presence of emission features and matrix effects [37] from the interferents and substrates greatly complicates the LIBS spectra.

Experimental

As part of a study funded by ARL (Battelle Study No. B027-G664451), the Battelle Eastern Science and Technology (BEST) Center Facility generated a library of laser-induced breakdown spectra using simulants for chemical and biological agents as well as selected interferents dispensed onto 1"×1" substrate coupons made of polycarbonate, stainless steel, and aluminum foil. The polycarbonate and stainless steel coupons had a protective coating that was removed immediately prior to use, while the aluminum foil coupons were cleaned with 70% isopropanol. Potential environmental interferents selected for this study included dolomitic limestone (Lime, NIST Standard Reference Materials, Catalog No. SRM 88b) and ovalbumin (Ova, Fisher Catalog No. BP2535-5). Table 1 lists the biological and chemical agent simulants, concentrations (in colony-forming units or plaque-forming units), and sources. Controls for the simulants included: Luria broth (Luria, Growcells Catalog No. MBLE-3030), inoculation control for MS-2 bacteriophage; 1× phosphate-buffered saline with 1% bovine serum albumin (BSA), inoculation control for α -hemolysin; and 1 M chloroform (MP Biomedical Catalog No. 194002), inoculation control for 2-chloroethyl ethyl sulfide (CEES). The substrate coupons were inoculated ten times with 10- μ L droplets of the simulant, control, or interferent. For mixtures, the coupon was first inoculated with the simulant or control, and then 100 μ L of the interferent was added directly to the simulant or control. The samples were dried overnight in a Biological Safety Cabinet prior to acquisition of the LIBS spectra.

The library of single-shot LIBS spectra was acquired by Battelle personnel using the LIBS Pelicase PL100-GEO instrument developed by A3 Technologies. The PL100-GEO instrument is a compact LIBS system with a 25-mJ laser (1,064 nm) and multi-channel CCD spectrometer with continuous wavelength coverage from 195 to 966 nm. Seventy-two unique residue/substrate combinations were obtained (Table 2). The library of spectra was then provided to ARL for analysis. PLS-DA analysis was performed using the PLS_Toolbox 5.0.3 software (Eigenvector Research, Inc.).

Results and discussion

Assignment of spectral features

The first step in the analysis of the data provided by Battelle was to identify the emission features in the LIBS spectra. Since the substrate emission lines were also present in the LIBS spectra of the residues, spectra of blank substrate coupons were used to determine the spectral

Table 1 Biological and chemical agent simulants

Biological/chemical agent	Description	Concentration	Source
<i>Bacillus atrophaeus</i> (BA) spores	Simulant of <i>Bacillus anthracis</i> (anthrax) spores	5.5×10^6 CFU/coupon	Internal Battelle source
<i>E. coli</i>	Simulant for Gram (-) bacterial biothreat agents	4.0×10^7 CFU/coupon	ATCC catalog no.15597
MS-2 bacteriophage (Phage)	Simulant of viral select agents such as <i>Variola virus</i> (smallpox)	2.8×10^{10} PFU/coupon	ATCC catalog no. 15597-B1
α -Hemolysin from <i>Staphylococcus aureus</i> (Hemo)	Simulant for biological toxins such as botulinum neurotoxin	0.1 mg/mL	Sigma catalog no. H9395
2-Chloroethyl ethyl sulfide (CEES)	Simulant for mustard agent	Diluted in chloroform to 250 mg/mL	Sigma catalog no. 242640
Dimethyl methylphosphonate (DMMP)	Simulant for the nerve agent sarin (GB)	Diluted in water to 250 mg/mL	Sigma catalog no. D169102

CFU colony-forming units, PFU plaque-forming units

contribution for each substrate. Fig. 1 shows typical emission spectra for the three substrates. While the polycarbonate and aluminum spectra contain relatively few strong emission lines, the stainless steel spectra consists of many strong emission lines, primarily due to Fe. The stainless steel spectra also contained emission lines from Cr, Mn, Ca, H, K, Na, O, and N. The O and N emission lines are primarily from the surrounding air, while Ca, Na, and K are common contaminants. Emission

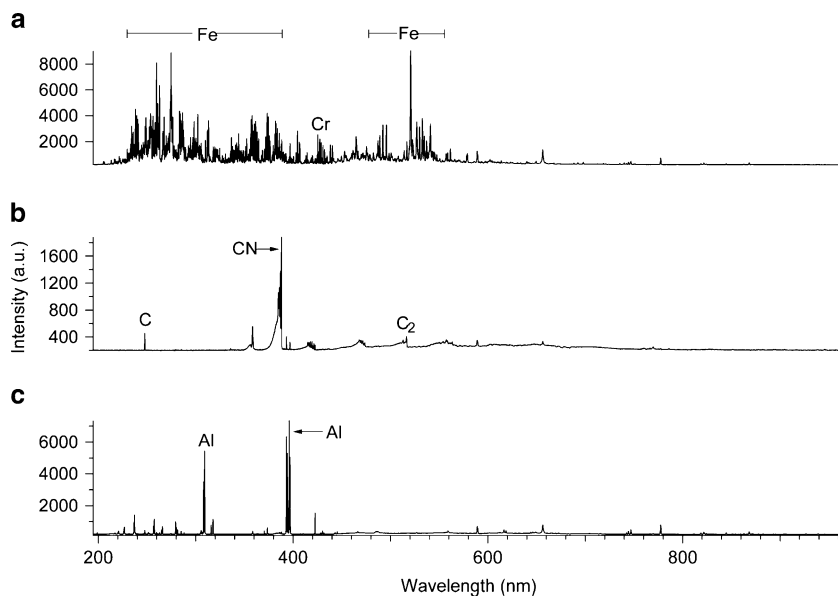
features in the polycarbonate spectra include C, C₂, CN, H, Ca, Na, and O. The C₂ emission is likely due to both contributions from ablated molecular fragments and recombination reactions; however, since polycarbonate does not contain any N (only C, H, and O), the CN emission is due to recombination reactions with atmospheric N [38]. The aluminum spectra contain Al, Mg, Si, Sr, C, CN, H, Ca, Na, K, and O emission lines. The presence of C, CN, and H in the aluminum spectra indicates that there was some organic

Table 2 Residue/substrate sample types (no. of single-shot LIBS spectra acquired listed in parentheses)

Aluminum substrate		Steel substrate		Polycarbonate substrate	
Pure residues	Mixtures	Pure residues	Mixtures	Pure residues	Mixtures
Blank (83)	Al+BA+Lime (30)	Blank (34)	Steel+BA+Lime (30)	Blank (25)	Poly+BA+Lime (30)
BA (30)	Al+BA+Ova (30)	BA (30)	Steel+BA+Ova (30)	BA (30)	Poly+BA+Ova (30)
CEES (30)	Al+CEES+Lime (30)	BSA (30)	Steel+BSA+Lime (30)	BSA (30)	Poly+BSA+Lime (30)
Chl (30)	Al+CEES+Ova (30)	CEES (30)	Steel+BSA+Ova (30)	CEES (33)	Poly+BSA+Ova (30)
Lime (30)	Al+Chl+Lime (30)	Chl (33)	Steel+CEES+Lime (28)	Chl (32)	Poly+CEES+Lime (33)
Ova (30)	Al+Chl+Ova (30)	DMMP (65)	Steel+CEES+Ova (36)	DMMP (79)	Poly+CEES+Ova (31)
		<i>E. coli</i> (30)	Steel+Chl+Lime (29)	<i>E. coli</i> (35)	Poly+Chl+Lime (28)
		Hemo (35)	Steel+Chl+Ova (31)	Hemo (35)	Poly+Chl+Ova (28)
		Lime (52)	Steel+DMMP+Lime (78)	Lime (29)	Poly+DMMP+Lime (37)
		Luria (35)	Steel+DMMP+Ova (91)	Luria (28)	Poly+DMMP+Ova (35)
		Ova (19)	Steel+ <i>E. coli</i> +Lime (30)	Ova (32)	Poly+ <i>E. coli</i> +Lime (31)
		Phage (29)	Steel+ <i>E. coli</i> +Ova (31)	Phage (28)	Poly+ <i>E. coli</i> +Ova (32)
			Steel+Hemo+Lime (30)		Poly+Hemo+Lime (35)
			Steel+Hemo+Ova (30)		Poly+Hemo+Ova (30)
			Steel+Luria+Lime (28)		Poly+Luria+Lime (30)
			Steel+Luria+Ova (32)		Poly+Luria+Ova (34)
			Steel+Phage+Lime (30)		Poly+Phage+Lime (28)
			Steel+Phage+Ova (29)		Poly+Phage+Ova (29)

Chl chloroform

Fig. 1 LIBS spectra of the **a** stainless steel, **b** polycarbonate, **c** aluminum foil substrate coupons. Prominent atomic and molecular emission features have been labeled



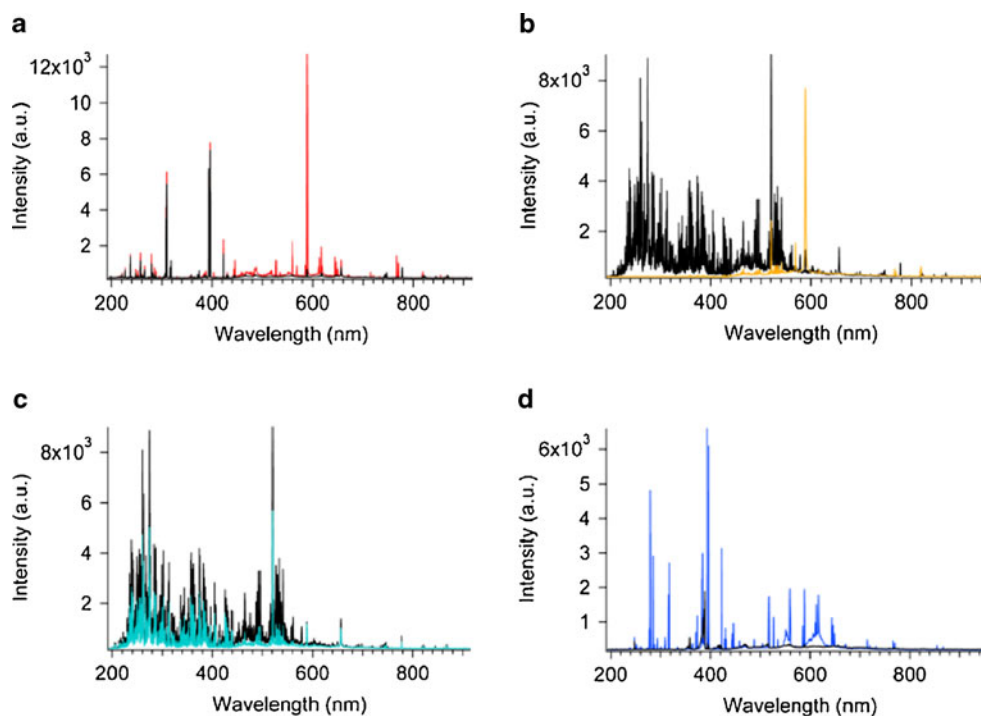
contamination on the foil, either from the manufacturing process or from the cleaning treatment with isopropanol.

A total of 188 emission lines from 20 atomic and molecular species were identified from the LIBS spectra of the residues based on published atomic [39] and molecular [40] databases. Fig. 2 shows some examples of residue spectra on different substrates. While some residues such as the α -hemolysin were relatively opaque to the 1,064-nm laser and suppressed the substrate emission lines, other residues such as the ovalbumin were transparent to the laser and resulted in strong substrate emission. The determination of whether emission lines originated from the residue

or the substrate was made based on the comparison of the residue spectra on different substrates. For example, Fig. 3 shows the spectra of dimethyl methylphosphonate (DMMP) on (a) steel and (b) polycarbonate. Since the steel contains Cr, it is not obvious from Fig. 3a that the DMMP residue contains Cr. By comparing the spectra of DMMP on polycarbonate to the blank polycarbonate, however, the presence of Cr in DMMP becomes evident.

Table 3 lists the species observed in the simulants, controls, and interferents. Singly ionized Ca, Fe, Mg, Mn, and Zn emission lines were observed in addition to emission from the neutral species. Although Sr I was not

Fig. 2 LIBS spectra of simulant residues on substrates (*black traces*): **a** BA (*red*) on aluminum, **b** Hemo (*orange*) on steel, and interferent residues **c** Ova (*aqua*) on steel, **d** Lime (*blue*) on polycarbonate



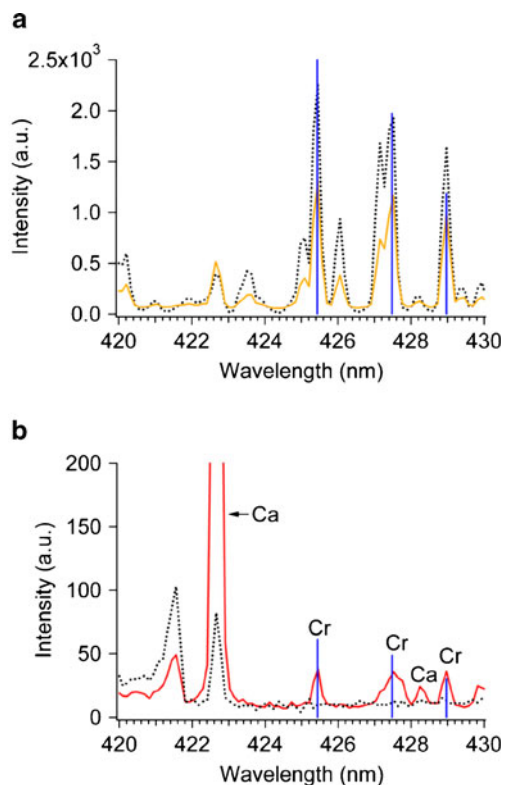


Fig. 3 LIBS spectra of **a** DMMP (solid/orange) on steel (dashed/black) and **b** DMMP (solid/red) on polycarbonate (dashed/black). The blue sticks represent the locations of the Cr emission lines according to the NIST database [39]

observed, the strong Sr II line at 407.77 nm was. CaOH emission at 554 nm and 622 nm was also identified in the limestone spectra and has been previously observed in other materials containing strong Ca lines [16, 27]. While the weak P emission was observed as expected in the DMMP and BSA (as well as the α -hemolysin, which contains the

BSA), no Cl or S emission was observed in the CEES spectra. Because the strongest transitions of Cl and S in the wavelength range of the spectrometer have upper levels that lie 10.4 and 7.87 eV above the ground state, they are difficult elements to detect with LIBS in solid materials under air [3], especially with the relatively low laser energy (25 mJ) used in this study. The presence of Al, Cr, Li, Mg, and Zn in the DMMP spectra suggests that the sample contains metallic contamination from an unknown source. Spectral features of the simulant residues were extremely difficult to discern in the presence of the interferents, as shown in Fig. 4. Emission features of the substrate, CEES residue, and Lime interferent are all present in the LIBS spectra of the mixture of CEES and Lime on steel (Fig. 4c).

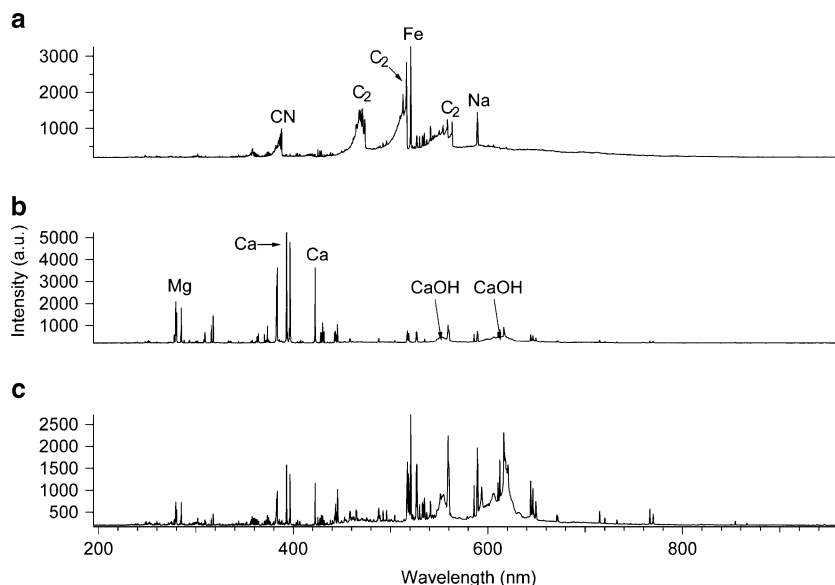
Chemometric model development

In order to classify the LIBS spectra of samples based on the simulant residues in the presence of interferents and different substrates, advanced chemometric analysis techniques were required. Several methods for building the PLS-DA model were developed and tested, as shown in Fig. 5. The first choice for building a PLS-DA model is the selection of input variables. The simplest choice is to use the full LIBS spectra. Based on our previous success discriminating explosive residues using pre-selected emission intensities and ratios, we also generated input variable data sets using the emission intensities and ratios of the simulant residues studied in this work. Models constructed using pure residue data from a single substrate were tested and compared with models based on multiple substrates. The independent test sets consisted of residue mixtures on the different substrates. The results from the two types of models (full-spectra vs. intensities/ratios) were fused

Table 3 Atomic and molecular emission species observed in LIBS spectra of the simulants, controls, and interferents

Hemo	BA	Phage	CEES	DMMP	<i>E. coli</i>	Chl	Luria	BSA	Lime	Ova
C	C	C	C	Al	Ca	C	C	C	Al	C
Ca	C ₂	Ca	C ₂	C	Cr	C ₂	Ca	Ca	C	Ca
H	Ca	K	Ca	Ca	Fe	Ca	K	H	Ca	CN
K	CN	Mg	CN	Cr	K	CN	Na	K	CaOH	H
Mg	H	Na	H	H	Na	H		Mg	CN	K
Na	K		Na	K		Na		Na	Fe	Na
P	Mg			Li				P	H	
Si	Mn			Mg					K	
	N			N					Li	
	Na			Na					Mg	
	O			O					N	
	Si			P					Na	
				Zn					O	
									Si	
									Sr	

Fig. 4 LIBS spectra of the residues **a** CEES, **b** Lime, and **c** CEES+Lime on the steel substrate



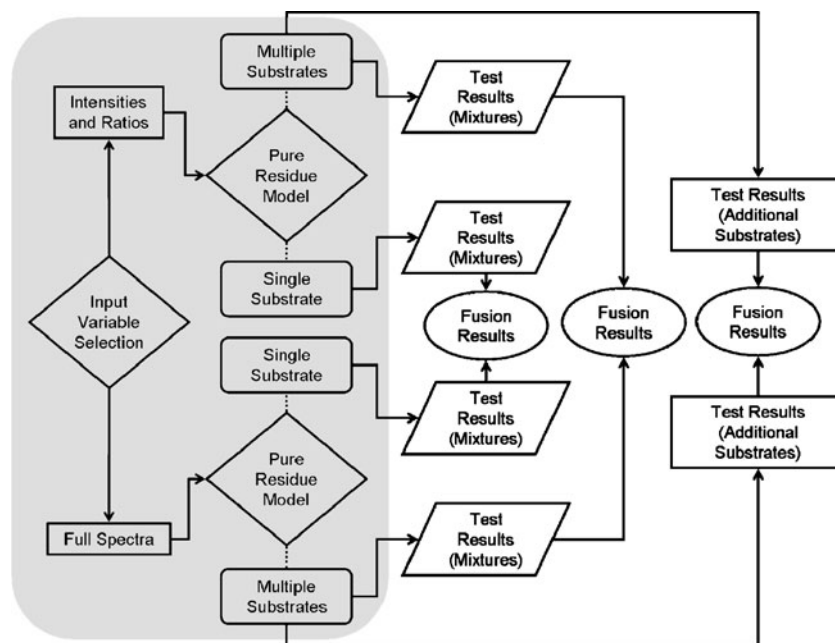
together to create a third type of model. The following sections describe the methods used to construct the models and the results obtained for each type of model.

Input variable selection

There are a number of advantages to using the full LIBS spectra as input variables for the PLS-DA model. Minimal preprocessing is required—for this work, the full spectra were normalized and mean-centered (i.e., the mean for all the spectra at each wavelength was subtracted from each individual spectrum so that the data at each wavelength represents the deviation from the average spectrum in the training data). In addition, a significant amount of variables

are available in the full spectra. The LIBS spectra from the PL100-GEO instrument contain 3,818 intensity channels, but broadband LIBS spectra from different types of spectrometers can contain tens of thousands of intensity channels. Significant computational power is therefore needed for large models containing many full broadband spectra. While the full spectra contain a considerable amount of relevant information, they also contain emission lines from the substrate and atmosphere, and many channels contain only baseline noise. Classification of the samples using full spectra is often based on matrix effects rather than the residues of interest. Thus, creating a robust model able to correctly classify spectra acquired under different experimental conditions is difficult [27, 41].

Fig. 5 Flowchart for the development (*shaded portion*) and testing of PLS-DA models to discriminate simulant residues on multiple substrates in the presence of interferents



By using selected emission intensities relevant to the samples of interest, emission lines due solely to substrates or interferences can be ignored. As in previous studies by our group, the background-corrected peak intensities of individual emission lines were combined to create summed intensities for each of the observed species (e.g., the summed H intensity was calculated by adding the intensities of the emission lines at 486 and 656 nm); the summed intensities were normalized to the total peak intensity of the observed species. Thirty summed, normalized emission intensities (including the 20 atomic and molecular species in Table 3 as well as individual summations of neutral and singly ionized species) were calculated and used to create 202 ratios, for total of 232 input variables. Simple ratios of the 30 summed intensities (e.g., Al/C, Mg I/Mg II, etc.) comprise 195 of the ratios; the remaining seven ratios are complex combinations of the summed intensities designed to provide additional information about the sample composition and the chemical reactions occurring in the laser-induced plasma [e.g., P/(C+H), (Ca+CaOH)/(O+H), etc.]. The use of summed intensity ratios reduces the effects of the shot-to-shot variability inherent in LIBS as well as the sample inhomogeneity and provides non-linear variables that provide additional sample discrimination based on composition [19]. Down-selecting the data in the LIBS spectra results in fewer input variables; thus, the necessary computational power for calculating large models is reduced. On the other hand, before building and testing the model, it must be decided which emission lines are important, and relevant variables may inadvertently be excluded. This step requires significant preprocessing by qualified personnel when the model is first being developed.

Once the two PLS-DA models based on full spectra and the down-selected variables were calculated for each sample set, the spectra from the test sets were tested against each model. The PLS-DA software calculates the predicted probabilities for each model class for each test spectrum. The predicted probabilities of each test spectrum obtained for the two models were multiplied together to create a “fusion model” that combines the advantages of using full spectra with the advantages of using intensities/ratios. In other words, for each sample spectrum tested against a model with N classes:

$$\begin{aligned} \text{Class 1 } P_{\text{fusion1}} &= P_{\text{full1}} \times P_{\text{int/ratio 1}} \\ \text{Class 2 } P_{\text{fusion2}} &= P_{\text{full2}} \times P_{\text{int/ratio 2}} \\ &\vdots \\ \text{Class N } P_{\text{fusion N}} &= P_{\text{fullN}} \times P_{\text{int/ratio N}}. \end{aligned}$$

This approach was previously used to improve the classification of explosive residues on multiple substrates [36].

Single-substrate models

Individual PLS-DA models were built based on the LIBS spectra of the pure residues (simulant, controls, and interferences) on each of the three substrates (aluminum, steel, and polycarbonate). Models based on full spectra and intensities/ratios were developed, and the results of the two models were fused to give a total of nine distinct single-substrate models. Cross-validation of the training set data was used to compare the ability of each model to correctly classify the pure residue samples on each substrate. Each single-shot sample spectrum was removed from the model one at a time and tested against the re-calculated model. Based on the Bayesian threshold automatically calculated for each class by the software in order to minimize the number of false positives and false negatives, each test spectrum was either: (1) correctly classified (i.e., the predicted Y value of the test spectrum was above the threshold for the correct class), (2) misclassified (i.e. the predicted Y value was below the threshold for the correct class but above the threshold for one of the other classes), or (3) unclassified (i.e., the predicted Y value is above the threshold for more than one class in the model or for none of the classes). An alternative method for cross-validation involves removing all of the single-shot spectra for a particular sample class (i.e., BA residue) and testing them against the re-calculated model; while the “leave one sample out” approach is believed to be more indicative of real-world performance in the case of sample materials that are relatively homogeneous and therefore produce similar LIBS spectra for each laser shot [42], the inhomogeneous nature of the residue samples results in significant shot-to-shot variation in the LIBS spectra (i.e., large intraclass variability). For this reason, the leave one shot out cross-validation method is sufficient for this application.

Five residues were applied to the aluminum substrate: BA, CEES, chloroform, limestone, and ovalbumin. PLS-DA models with six classes (including the substrate) were generated using the full-spectra and intensity/ratio input variables. The full-spectra models resulted in a higher correct classification rate than the intensity/ratio model (98.9% vs. 89.2%) with much lower misclassification and unclassified rates (Table 4). The fusion model results in similar classification performance for pure residues as the full-spectra models.

Spectra of the 11 pure residues (Table 3) were acquired on the steel substrate, and 12-class PLS-DA models were constructed. The steel substrate contains many more emission lines (see Fig. 1a), so the correct classification rate was only 84.5% for the full-spectra models and 55.3% for the intensity/ratio model. The most difficult residues to discriminate in the full-spectra models were the α -hemolysin (40.0%) and its control, BSA (53.3%). This

Table 4 Single-substrate model classification results (from cross-validation)

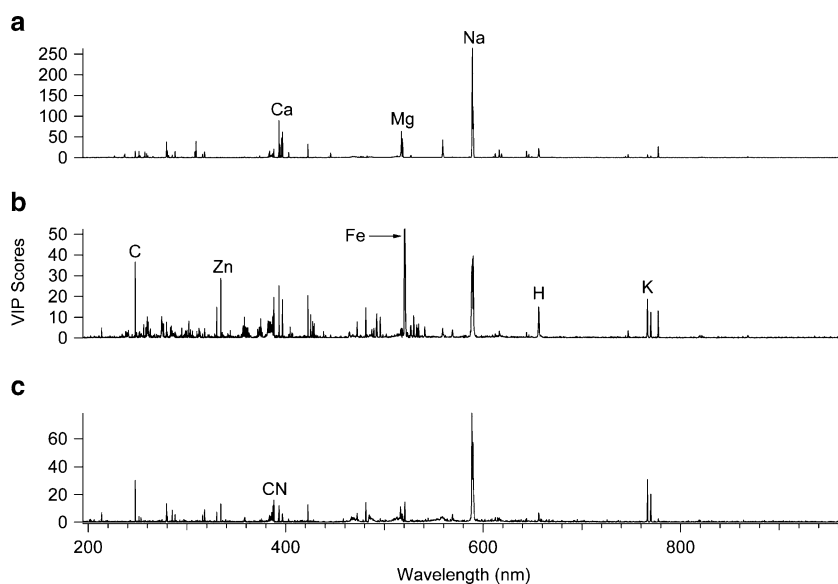
Model	Pure residues		
	Correct classification	Misclassified	Unclassified
Aluminum (full)	98.9%	0.0%	1.1%
Aluminum (int/ratios)	89.2%	0.8%	10.0%
Aluminum (fusion)	98.5%	0.0%	1.5%
Steel (full)	84.5%	0.7%	14.8%
Steel (int/ratios)	55.3%	7.1%	37.7%
Steel (fusion)	84.8%	0.5%	14.8%
Polycarbonate (full)	82.1%	0.0%	17.9%
Polycarbonate (int/ratios)	48.1%	4.5%	47.4%
Polycarbonate (fusion)	85.6%	1.1%	13.4%

indicates that most of the emission features in the α -hemolysin spectra are due to the BSA (the LIBS spectra of the two residues are nearly indistinguishable by eye). In addition to poor discrimination of the α -hemolysin (28.6%) and BSA (36.7%), the intensity/ratio model was unable to correctly classify the ovalbumin (15.8%) or bacteriophage (13.8%). Both residues have relatively few emission features which are mostly obscured by the steel emission spectra (Fig. 2b, c). As with the aluminum substrate, the steel fusion model gives results similar to the full-spectra models (Table 4).

PLS-DA models with 12-classes were also constructed for the 11 pure residues on the polycarbonate substrate. Unlike the aluminum and steel substrates, the polycarbonate emission spectra share a number of emission lines with the residues (e.g., C, CN, C₂). As with the steel substrate, the α -hemolysin and BSA residues were nearly indistinguishable on the polycarbonate; the full-spectra correct classification increased to 93.0% when the α -hemolysin and BSA are excluded. The intensity/ratio model also had

difficulty classifying the Luria broth (21.4%) and bacteriophage (17.9%) residues. While the full-spectra and intensity/ratio polycarbonate models gave comparable results to the steel substrate, the fusion model for the polycarbonate does not result in a significant improvement in the classification rate (Table 4).

Discrimination of the residue classes in the single-substrate models depends not only on the emission lines for each residue, but also on the substrate emission lines. A comparison of the VIP scores for the ovalbumin classes in each of the three substrate models (Fig. 6) shows that each full spectrum model depends not only on the ovalbumin emission lines for classification, but also on substrate emission lines (e.g., Fe) and emission lines present in other residues in the model (e.g., Mg, Zn). For this reason, the single-substrate models are not robust enough to correctly classify the residues on substrates not included in the model, and this approach would not work for applications where an unknown residue could be on any surface. By using a common swipe material to collect the residue,

Fig. 6 VIP Scores for the ovalbumin class in the single-substrate models for **a** aluminum, **b** steel, and **c** polycarbonate

however, a PLS-DA model could be built using the swipe material as a single substrate [15, 43, 44].

The ability of the models in Table 4 to correctly classify more complex samples outside the training sets was tested with the residue mixture data. The following procedure was used to identify the components of the residue mixtures using the PLS-DA models developed with the pure residue spectra. For each spectrum tested against a model, the highest predicted probability was used to determine the classification of the sample. Since multiple replicate single-shot spectra of each sample type (i.e., residue mixture) were acquired, the two classes with the most positive (unique) identifications for a particular sample were identified as the two components of the mixture. Table 5 shows the results from testing the residue mixtures on aluminum against the full-spectra and intensity/ratio models. While the full-spectrum model has difficulty identifying the CEES and chloroform residues in the limestone mixtures, it does pick up on the chloroform used to dilute the CEES. Only the intensity/ratio model is able to pick up all of the simulant and control residues in the presence of the interferents.

LIBS spectra of 18 residue mixtures on the steel and polycarbonate (Table 2) were tested against the steel and polycarbonate substrate models (using both full-spectra and intensity/ratios). Fusion of the predicted probabilities for the two types of models did not significantly improve the mixture results for any of the single-substrate models. Table 6 shows the truth–response table generated by applying the classification method described above for the aluminum mixtures to the residue mixtures on steel with the full-spectrum model. More than 75% of the simulant and control residues were correctly identified in the mixtures despite the extremely complex steel background (only ~60% were correctly identified by the polycarbonate models). The polycarbonate substrate shares a number of important carbon-related emission features with the residues, making discrimination of the residue mixtures on polycarbonate difficult. All of the steel and polycarbonate models had difficulty identifying the CEES, chloroform, DMMP, and *Escherichia coli* in the presence of the limestone interferent. While the models picked up on the

interferent in the mixture 40–75% of the time (depending on input variable selection and substrate), they also identified additional components in the sample, such as the substrate, solvent, or growth medium.

For the full-spectrum polycarbonate model, the chloroform was identified in the CEES+Ova mixture (as it was for the full-spectrum aluminum substrate model). The BSA was identified by both the full-spectra (Table 6) and intensity/ratio steel models in the α -hemolysin mixtures (three out of four mixtures). Because BSA is present in the α -hemolysin, this result is actually a correct classification. Unfortunately, many of the emission features used by the models to identify some of the threat simulants are due to the solvents or growth media (controls). Consequently, testing of the models with the residue mixtures resulted in several false positives, e.g. BSA mixtures were sometimes classified as α -hemolysin (six out of eight mixtures on steel and polycarbonate), and the Luria broth was sometimes classified as bacteriophage (seven out of eight mixtures on steel and polycarbonate). This is an issue since the presence of certain solvents or growth media, while perhaps suspicious in some contexts, does not necessarily imply that chemical or biological threats are also present.

The Luria broth was provided by Battelle as a control for the bacteriophage simulant; however, in addition to picking up the Luria broth in the bacteriophage mixtures, many of the steel and polycarbonate models also classified the *E. coli* (five out of eight mixtures) and BA (three out of eight mixtures) residues with the Luria broth. The Luria broth used as a control for this study contains tryptone, yeast extract, and NaCl. According to the vendor Website, the *E. coli* spores were grown in ATCC medium 271, which also contains tryptone, yeast extract, and NaCl (as well as glucose, CaCl₂, and thiamine). The BA spores were grown in trypticase soy broth, which contains tryptone and NaCl (as well as soytone, dextrose, and dipotassium phosphate). Therefore, the classification of the BA and *E. coli* with the Luria broth is reasonable. Several groups have previously shown that LIBS can be used to identify the growth media for bacterial spores [12, 17].

Table 5 PLS-DA model results for the residue mixtures on the aluminum substrate (the percentage of uniquely identified single-shot spectra classified with each residue type is listed in parentheses)

Test sample	Full-spectra model		Intensity/ratio model	
	1 st (%)	2 nd (%)	1 st (%)	2 nd (%)
Al+BA+Lime	BA (40)	Lime (17)	BA (63)	Lime (27)
Al+BA+Ova	BA (80)	Ova (20)	BA (73)	Ova (10)
Al+CEES+Lime	Lime (40)	Chl (7)	Lime (40)	CEES (10)
Al+CEES+Ova	CEES (53)	Chl (13)	Lime (20)	CEES (17)
Al+Chl+Lime	Lime (73)	–	Lime (77)	Chl (20)
Al+Chl+Ova	Chl (47)	Ova (33)	Chl (57)	Ova (27)

Table 6 PLS-DA model results for the residue mixtures on the steel substrate using the full-spectra (color key: red (R)=correct simulat/control identification, blue (B)=correct interferent identification, green (G)=correct component identification, black (+)=false positive, gray (-)=false negative)

TRUTH	RESPONSE											
	BA	BSA	CEES	Chl	DMMP	Ecoli	Hemo	Luria	Phage	Lime	Ova	Steel
Steel+BA+Lime [21]	R									B		
Steel+BA+Ova [26]	R							G				
Steel+BSA+Lime [30]		R					+			B		
Steel+BSA+Ova [30]		R					+					
Steel+CEES+Lime [9]			-							B		
Steel+CEES+Ova [33]			R					+			B	
Steel+Cl+Lime [5]	+									B		
Steel+Cl+Ova [31]				R								
Steel+DMMP+Lime [78]					-					B		
Steel+DMMP+Ova [76]					R						B	
Steel+Ecoli+Lime [13]						-		G				
Steel+Ecoli+Ova [31]						R		G				
Steel+Hemo+Lime [30]							R					
Steel+Hemo+Ova [30]							R					
Steel+Luria+Lime [27]								R	+			
Steel+Luria+Ova [32]								R	+			
Steel+Phage+Lime [30]								G	R			
Steel+Phage+Ova [29]								G	R			

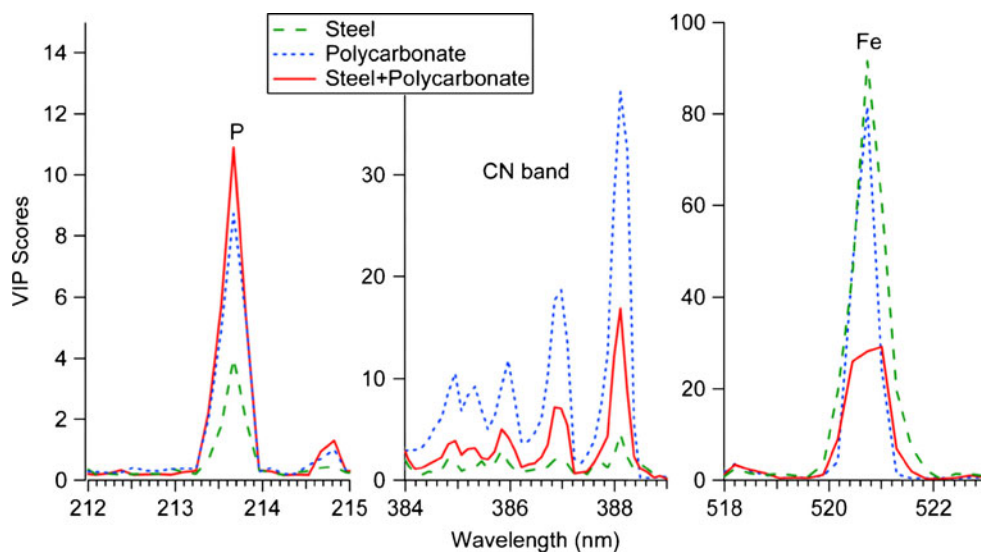
The number of uniquely identified single-shot spectra for each residue mixture is listed in brackets

Multiple-substrate models

As a first step toward generating a more robust PLS-DA model capable of classifying chemical and biological residues on multiple substrates, we created a model in which each residue class contains spectra of the residue on both the steel and polycarbonate substrates. This approach is analogous to our previous work involving the detection of explosive residues on multiple substrates [36]. The goal

was to reduce the dependence of the model on the substrate emission features and related matrix effects by training the model to focus on the residue emission lines. Fig. 7 shows the VIP scores for the α -hemolysin class in the steel (green) and polycarbonate (blue) single-substrate models and the multiple-substrate model (red). The P line from the α -hemolysin residue contributes the most to the multiple-substrate model (compared with the single-substrate models). On the other hand, the CN and Fe emission, which are due to

Fig. 7 Selected regions of the VIP scores for the α -hemolysin class in the steel and polycarbonate single-substrate models and the multiple substrate model (steel and polycarbonate). The P emission is due to the α -hemolysin residue, while the CN and Fe emission is primarily due to the polycarbonate and steel substrates, respectively



the polycarbonate and steel substrates, respectively, contribute most strongly to the α -hemolysin discrimination in the single-substrate models.

The cross-validation results for the multiple-substrate model (Table 7) were similar to those obtained using the intensity/ratio single-substrate models (Table 4). The biggest difference was that the full-spectra models were much more accurate when the model was constructed using residues on a single substrate only. This provides additional evidence that correct classification in full-spectra models depends on substrate emission lines as well as the relevant residue emission lines. Unlike the single-substrate models, fusion of the results from the full-spectra and intensity/ratio models improves the classification results for the multiple-substrate model (from 47.4% to 63.1%). In addition to having a higher correct classification rate, the misclassified rate is the lowest for the fusion results, since incorrect identifications in one model can be offset by the correct identification in the other model [36]. Most of the unclassified spectra in Table 7 result from the classification of pure residues with multiple classes (e.g., substrate and residue, threat agent and control, spore and growth medium, etc.). For example, the steel+Ova test spectra often classify with both the steel and ovalbumin classes in the steel/polycarbonate fusion model, and the Hemo residues on both steel and polycarbonate classify with both BSA and Hemo.

The results of testing the residue mixture spectra on the polycarbonate and steel substrates against the multiple-substrate models using both the full-spectra and intensity/ratios were very similar to the results obtained from the single-substrate models (not shown). In order to gauge the ability of the multiple-substrate models to correctly classify residues on a third substrate not included in the training set, the LIBS spectra of the pure residues on the aluminum substrate were tested against the models built using the polycarbonate and steel substrate spectra. Unfortunately, the results were not encouraging. While the BA (17%), chloroform (10%), and limestone (23%) residues were correctly classified with the full-spectra model, only the limestone residue was correctly classified with the intensity/ratio (10%) and fusion (67%) models. The full-spectra model classified most of the residues on aluminum as DMMP (63–100% for residues other than limestone, resulting in false positives for a chemical threat simulant),

while the intensity/ratio and fusion models resulted in random misclassifications (e.g., BA on aluminum classified with BSA 80% of the time in the fusion model). The residues on the aluminum substrate likely classified with the DMMP due to the presence of Al emission features in the DMMP. Similarly, the multiple-substrate models were able to correctly identify limestone in the residue mixtures on aluminum, but had few other correct classifications. Clearly, the models consisting of only two substrates are not robust enough to handle additional substrates.

Conclusions

We have shown that, with careful consideration of the input variable selection and model-building, PLS-DA models capable of correctly classifying chemical and biological residues in mixtures containing interferents can be developed. In addition to identifying the major components of such residue mixtures, minor components such as growth media and solvents can be identified (assuming the model has been trained appropriately). We find that, while the full-spectra models based on single substrates give the best results overall, the intensity/ratio models help differentiate certain residues with fewer spectral features (e.g., CEES, DMMP, *E. coli*). The full-spectra models also strongly depend on substrate emission lines and matrix effects for correct classification of test residues. Fusion of the full-spectra and intensity/ratio models result in only marginally improved correct classification for the single substrate models, but a significant improvement for the multiple-substrate models.

While classification of the residue mixtures on complex substrate backgrounds works surprisingly well, the controls and simulants were confused by some of the models (e.g., BSA/Hemo and Luria/*E. coli*), especially in the presence of interferents. Based on our experience building models for explosive residue detection, we expect that increasing the number of classes in the models (i.e., the types of residues) will result in improved classification. Also, the combined polycarbonate and steel substrate model could not correctly classify residues on aluminum. Increasing the number of substrates used for the model training set should enable the model to correctly classify residues on a wider range of substrates. Ultimately, the best solution for residue detec-

Table 7 Multiple substrate model classification results (from cross-validation)

Model	Correct classification	Misclassified	Unclassified
Full-spectra	54.8%	2.0%	43.3%
Intensity/ratios	47.4%	3.3%	49.3%
Fusion results	63.1%	1.8%	35.1%

tion (especially in cases where standoff detection is not required) may be to use a common swipe material to sample the residue, with PLS-DA models specifically designed to classify residues on the swipe material.

Acknowledgments Sample preparation and data collection were performed by Battelle technicians Dana Short, Tyler Goralski, and Stephanie McCaslin (study director, Kimberly Weber) with funding from the US Army Research Laboratory. LIBS spectra were collected by Battelle at A3 Technologies with the assistance of Diane Wong. The author also wishes to thank Richard Rossman (Battelle) for providing the spectral data and experimental details and Frank De Lucia, Jr., for a critical reading of the manuscript.

References

1. Cremers DA, Radziemski LJ (2006) Handbook of Laser-Induced Breakdown Spectroscopy. John Wiley & Sons, Ltd., West Sussex, England
2. Pasquini C, Cortez J, Silva LMC, Gonzaga FB (2007) Laser Induced Breakdown Spectroscopy. J Braz Chem Soc 18(3):463–512
3. Dudragne L, Adam P, Amouroux J (1998) Time-resolved laser-induced breakdown spectroscopy: Application for qualitative and quantitative detection of fluorine, chlorine, sulfur, and carbon in air. Appl Spectrosc 52(10):1321–1327
4. DeLucia FC Jr, Samuels AC, Harmon RS, Walters RA, McNesby KL, LaPointe A, Winkler RJ Jr, Miziolek AW (2005) Laser-induced breakdown spectroscopy (LIBS): a promising versatile chemical sensor technology for hazardous material detection. IEEE Sens J 5(4):681–689
5. Stearns JA, McElman SE, Dodd JA (2010) Identification of vapor-phase chemical warfare agent simulants and rocket fuels using laser-induced breakdown spectroscopy. Appl Opt 49(13):C8–C15
6. Morel S, Leone N, Adam P, Amouroux J (2003) Detection of bacteria by time-resolved laser-induced breakdown spectroscopy. Appl Opt 42(30):6184–6191
7. Samuels AC, DeLucia FC Jr, McNesby KL, Miziolek AW (2003) Laser-induced breakdown spectroscopy of bacterial spores, molds, pollens, and protein: initial studies of discrimination potential. Appl Opt 42(30):6205–6209
8. Kim DS, Oh B, Lee H (2004) Effect of liquid film on near-threshold laser ablation of a solid surface. Appl Surf Sci 222(1–4):138–147
9. Munson CA, De Lucia Jr FC, Piehler T, McNesby KL, Miziolek AW (2005) Investigation of statistics strategies for improving the discriminating power of laser-induced breakdown spectroscopy for chemical and biological warfare agent simulants. Spectrochim Acta B 60(7–8):1217–1224
10. Baudelet M, Guyon L, Yu J, Wolf JP, Amodeo T, Frejafon E, Laloi P (2006) Femtosecond time-resolved laser-induced breakdown spectroscopy for detection and identification of bacteria: A comparison to the nanosecond regime. J Appl Phys 99(8):84701
11. Diedrich J, Rehse SJ, Palchadhuri S (2007) *Escherichia coli* identification and strain discrimination using nanosecond laser-induced breakdown spectroscopy. Appl Phys Lett 90:163901
12. Diedrich J, Rehse SJ, Palchadhuri S (2007) Pathogenic *Escherichia coli* strain discrimination using laser-induced breakdown spectroscopy. J Appl Phys 102 (1):014702/1-014702/8
13. Rehse SJ, Diedrich J, Palchadhuri S (2007) Identification and discrimination of *Pseudomonas aeruginosa* bacteria grown in blood and bile by laser-induced breakdown spectroscopy. Spectrochimica Acta, Part B 62(10):1169–1176
14. Snyder EG, Munson CA, Gottfried JL, De Lucia Jr FC, Gullett B, Miziolek A (2008) Laser-induced breakdown spectroscopy for the classification of unknown powders. Appl Opt 47(31):G80–G87
15. Munson CA, Gottfried JL, Snyder EG, De Lucia FC Jr, Gullett B, Miziolek AW (2008) Detection of indoor biological hazards using the man-portable laser induced breakdown spectrometer. Appl Opt 47(31):G48–G57
16. Gottfried JL, De Lucia Jr FC, Munson CA, Miziolek AW (2008) Standoff Detection of Chemical and Biological Threats Using Laser-Induced Breakdown Spectroscopy. Appl Spectrosc 62 (4):353–363
17. De Lucia Jr FC, Gottfried JL, Munson CA, Miziolek AW (2009) Current Status of Standoff LIBS Security Applications at the United States Army Research Laboratory. Spectrosc 24(6):1
18. Rehse SJ, Mohaidat QI, Palchadhuri S (2010) Towards the clinical application of laser-induced breakdown spectroscopy for rapid pathogen diagnosis: the effect of mixed cultures and sample dilution on bacterial identification. Appl Opt 49(13):C27–C35
19. Gottfried JL, De Lucia Jr FC, Munson CA, Miziolek AW (2009) Laser-induced breakdown spectroscopy for detection of explosives residues: a review of recent advances, challenges, and future prospects. Anal Bioanal Chem 395:283–300
20. Boyain-Goitia AR, Beddows DCS, Griffiths BC, Telle HH (2003) Single-pollen analysis by laser-induced breakdown spectroscopy and Raman microscopy. Appl Opt 42(30):6119–6132
21. Hybl JD, Lithgow GA, Buckley SG (2003) Laser-Induced Breakdown Spectroscopy Detection and Classification of Biological Aerosols. Appl Spectrosc 57(10):1207–1215
22. Gibb-Snyder E, Gullett B, Ryan S, Oudejans L, Touati A (2006) Development of size-selective sampling of *Bacillus anthracis* surrogate spores from simulated building air intake mixtures for analysis via laser-induced breakdown spectroscopy. Appl Spectrosc 60(8):860–870
23. Hybl JD, Tysk SM, Berry SR, Jordan MP (2006) Laser-induced fluorescence-cued, laser-induced breakdown spectroscopy biological-agent detection. Appl Opt 45(34):8806–8814
24. Sirven J-B, Salle B, Mauchien P, Lacour J-L, Maurice S, Manhes G (2007) Feasibility study of rock identification at the surface of Mars by remote laser-induced breakdown spectroscopy and three chemometric methods. J Anal At Spectrom 22:1471–1480
25. Clegg SM, Sklute E, Dyar MD, Barefield JE, Wiens RC (2009) Multivariate analysis of remote laser-induced breakdown spectroscopy spectra using partial least squares, principal component analysis, and related techniques. Spectrochim Acta B 64(1):79–88
26. Doucet FR, Faustino PJ, Sabsabi M, Lyon RC (2008) Quantitative molecular analysis with molecular bands emission using laser-induced breakdown spectroscopy and chemometrics. J Anal At Spectrom 23(5):694–701
27. Gottfried JL, Harmon RS, De Lucia Jr FC, Miziolek AW (2009) Multivariate Analysis of LIBS Spectra for Geomaterial Classification. Spectrochimica Acta Part B 64:1009–1019
28. Gottfried JL, De Lucia Jr FC, Munson CA, Miziolek AW (2008) Strategies for residue explosives detection using laser-induced breakdown spectroscopy. J Anal At Spectrom 23:205–216
29. Labbé N, Swamidoss IM, André N, Martin MZ, Young TM, Rials TG (2008) Extraction of information from laser-induced breakdown spectroscopy spectral data by multivariate analysis. Appl Opt 47(31):G158–G165
30. Rehse SJ, Mohaidat QI (2009) The effect of sequential dual-gas testing on laser-induced breakdown spectroscopy-based discrimination: Application to brass samples and bacterial strains. Spectrochim Acta B 64(10):1020–1027
31. Barker M, Rayens W (2003) Partial least squares for discrimination. J Chemometrics 17:166–173

32. Bylesjo M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J (2006) OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemometrics* 20(8–10):341–351
33. Remus JJ, Gottfried JL, Harmon RS, Draucker A, Baron D, Yohe R (2010) Archaeological applications of laser-induced breakdown spectroscopy: an example from the Coso Volcanic Field, California, using advanced statistical signal processing analysis. *Appl Opt* 49(13):C120–C131
34. Wise BM, Gallagher NB, Bro R, Shaver JM, Windig W, Koch RS (2006) PLS_Toolbox Version 4.2 Reference Manual. Eigenvector Research, Inc., Wenatchee, WA
35. Chong I-G, Jun C-H (2005) Performance of some variable selection methods when multicollinearity is present. *Chemom Intell Lab Syst* 78(1–2):103–112
36. Gottfried JL, De Lucia Jr FC, Miziolek AW (2009) Discrimination of explosive residues on organic and inorganic substrates using laser-induced breakdown spectroscopy. *J Anal At Spectrom* 24(3):288–296
37. Windom BC, Hahn DW (2009) Laser ablation—laser induced breakdown spectroscopy (LA-LIBS): A means for overcoming matrix effects leading to improved analyte response. *J Anal At Spectrom* 24:1665–1675
38. Park HS, Nam SH, Park SM (2005) Time-resolved optical emission studies on the laser ablation of a graphite target: The effects of ambient gases. *J Appl Phys* 97(11):113103/1–113103/5
39. Ralchenko Y, Kramida AE, Reader J, Team NA (2008) NIST Atomic Spectra Database (version 4.0). National Institute of Standards and Technology. <http://physics.nist.gov/asd3>. Accessed 18 June 2010
40. Pearse RWB, Gaydon AG (1976) *The Identification of Molecular Spectra*, 4th edn. John Wiley & Sons, Inc, New York
41. De Lucia Jr. FC, Gottfried JL (2011) Influence of variable selection on partial least squares discriminant analysis models for explosive residue discrimination. *Spectrochimica Acta Part B* in press (doi:10.1016/j.sab.2010.12.007)
42. Alvey DC, Morton K, Harmon RS, Gottfried JL, Remus JJ, Collins LM, Wise MA (2010) Laser-induced breakdown spectroscopy-based geochemical fingerprinting for the rapid analysis and discrimination of minerals: the example of garnet. *Appl Opt* 49(13):C168–C180
43. Chinni R, Cremers DA, Multari R (2010) Analysis of material collected on swipes using laser-induced breakdown spectroscopy. *Appl Opt* 49(13):C143–C152
44. Gottfried JL (2010) Comparison of Swipe Materials for Chemical, Biological, and Explosive Residue Discrimination with Laser-Induced Breakdown Spectroscopy (US Army Research Laboratory, Aberdeen Proving Ground, MD), ARL-TR-5231